Exploratory Analysis:
Netflix Movies and TV Shows

Harini Rao | Neelam Baxi | Veronica Ratcliff | Zen Acar

Data Analysis & Visualization
Project 1
3/24/20

Motivation and Summary

We find ourselves turning on our TVs and selecting the Netflix app on a daily basis, especially now that we are all home more than ever. We were interested in exploring a Netflix dataset to see what content Netflix has added over the past 11 years (2008–2019). Due to the spread of COVID-19, everyone across the Globe is either working from home or trying to entertain their kids and the best source of entertainment is TV/Internet.

Dataset source used:
Netflix – Kaggle: https://www.kaggle.com/shivamb/netflix-shows
Female Directors List – IMDB: https://www.imdb.com/list/ls027192075/

The Netflix dataset consists of tv shows and movies available on Netflix as of January 2020. The dataset contains 6234 titles (TV Series and Movies) available on Netflix.

The Female Director List consists of 197 rows, with Female Directors and Titles of Movies and other unnecessary columns which we removed for our analysis.

Our group wanted to take a deep dive into Netflix and explore what content it offers and which countries have produced that content. Additionally, as an all women team, we were interested in exploring how Female Directors are represented in terms of content on Netflix. As we explored, more questions arose.

Questions/Major Findings from Data Analysis

Distribution of Content:

1. How many Movies/TV Shows have been added on Netflix in the past 11 years?

Production Type Distribution

Content added over the years


TV Shows & Movies Added By Year
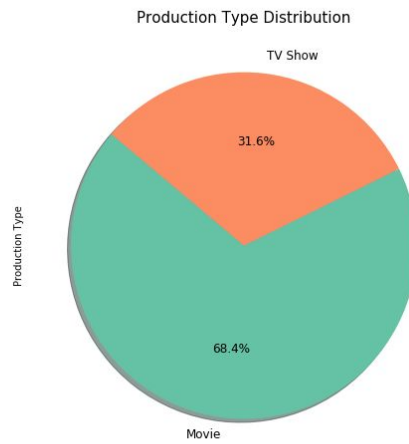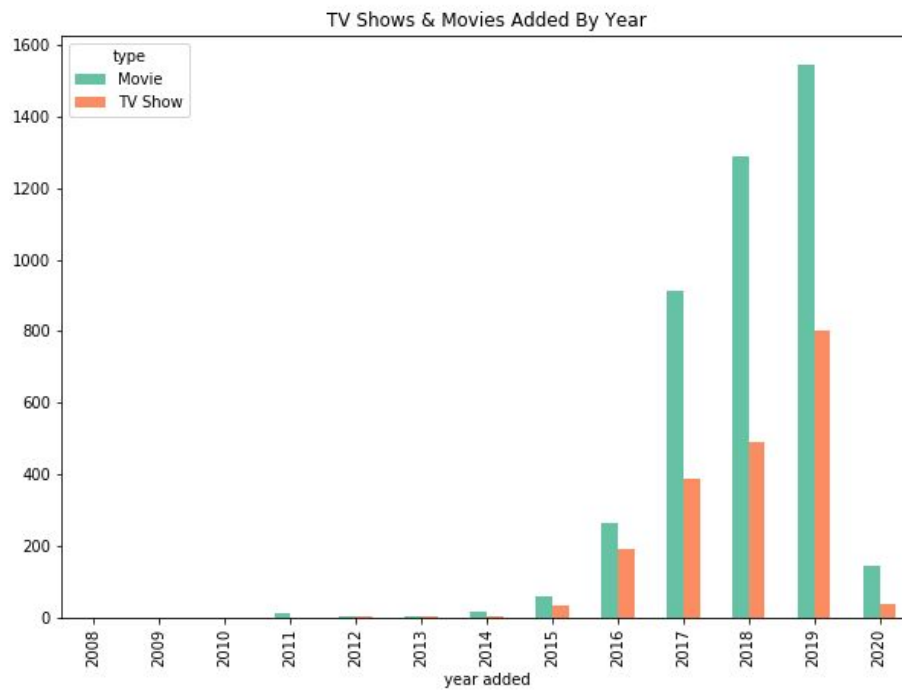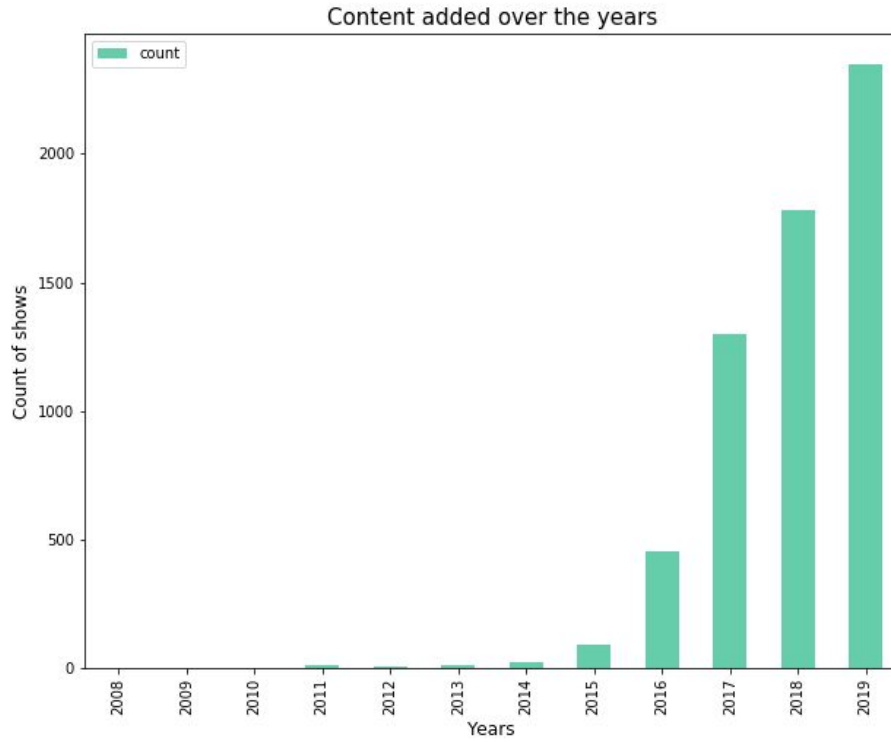
Between 2008-January 2020, of the available content, two-thirds (or slightly above, 68.4%) of it is movies and TV shows make up the remaining one-third (31.6%). Every year there is an increase in the amount of movies and TV shows added to the Netflix library, with 2019 having the most content added (2019: 1546 movies and 803 TV shows). There was a significant increase from 2015-2019.

Growth of content added over the years by Type
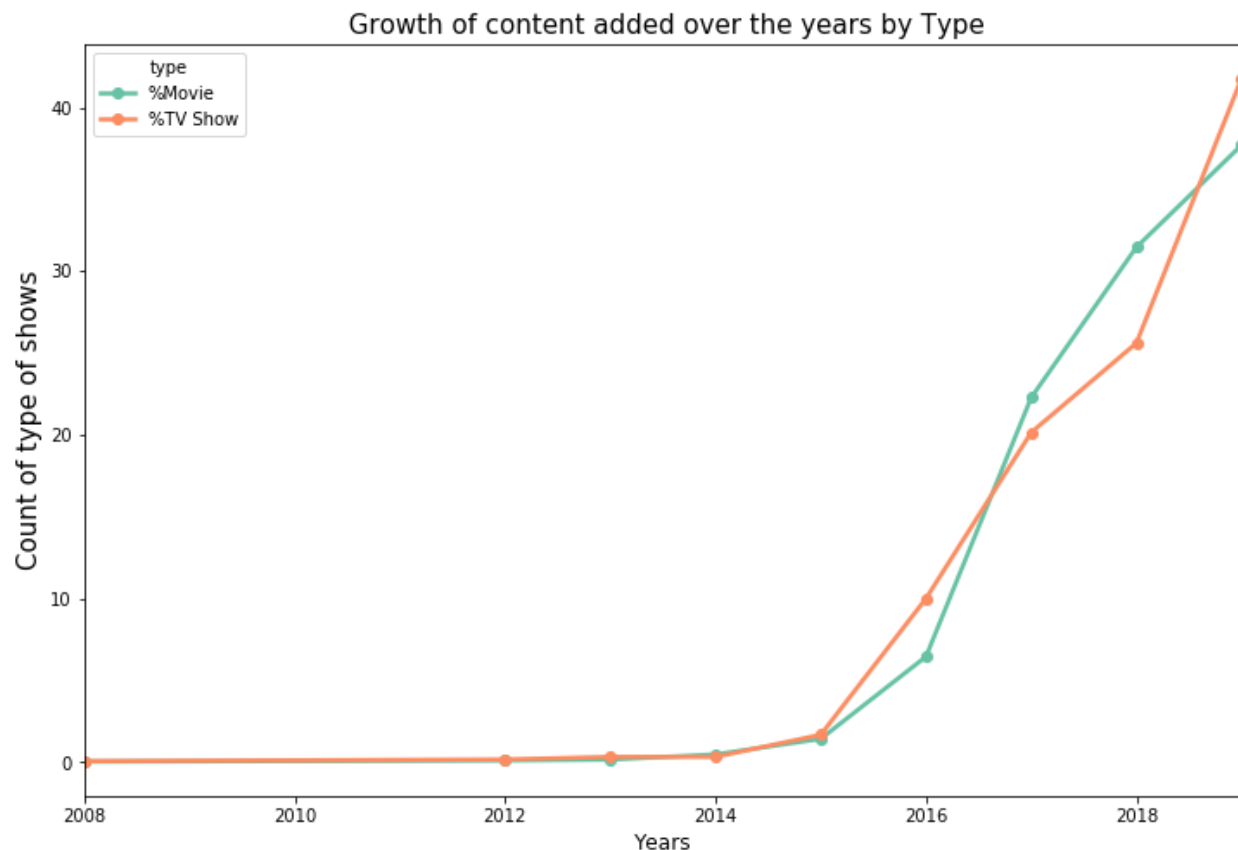
The above plot indicates that more movies seem to be added to Netflix content than that of TV shows from 2008-2018. This suggests that members may be preferring TV shows versus movies. However, in the past two years, there has been a shift in the percentage of content type being added. From 2018, a significant growth has been observed in the percentage of TV shows added as opposed to Movies.

What is the correlation between Movie Release Year and Year Added on Netflix?
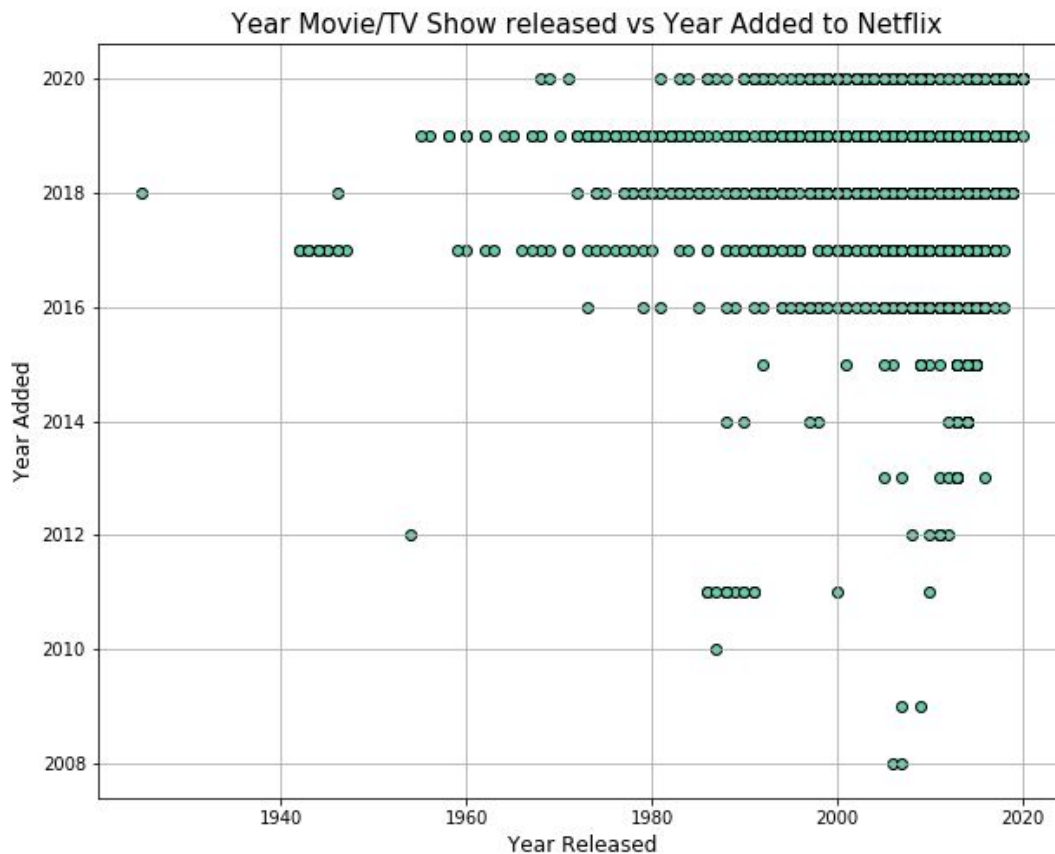

The correlation between Release Year and Year Added to Netflix is 0.02.

The above plot analyzes the relationship between the year a movie was released and when it was added to Netflix. There is less to no correlation between the two variables. We do see that some of the movies that were released between 1980 to 2000 weren't added to Netflix until 2011. We also see one movie/TV show that was released in 2020 and added in 2020.

2. What countries have produced content that is offered on Netflix?


Content added by Country
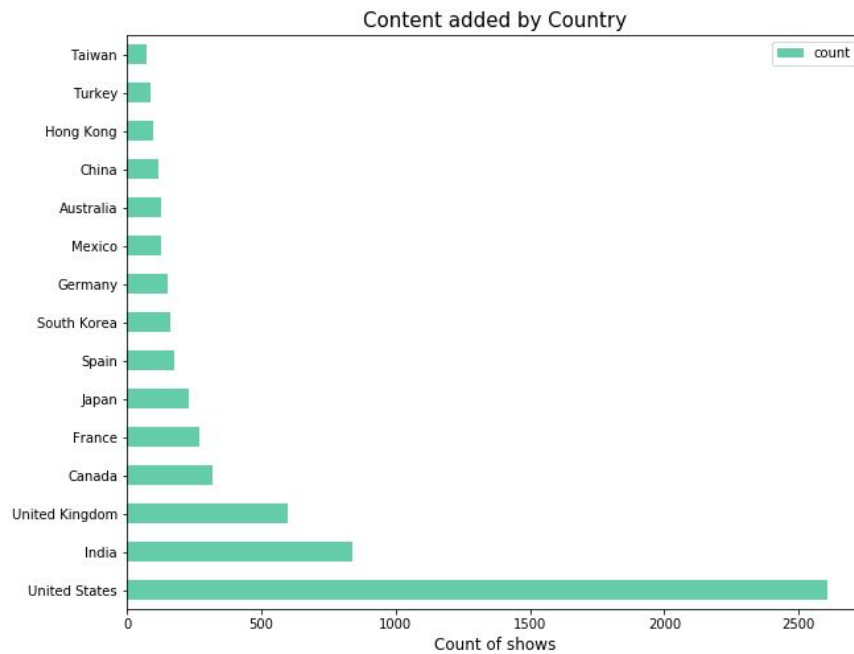
Geographical distribution of content
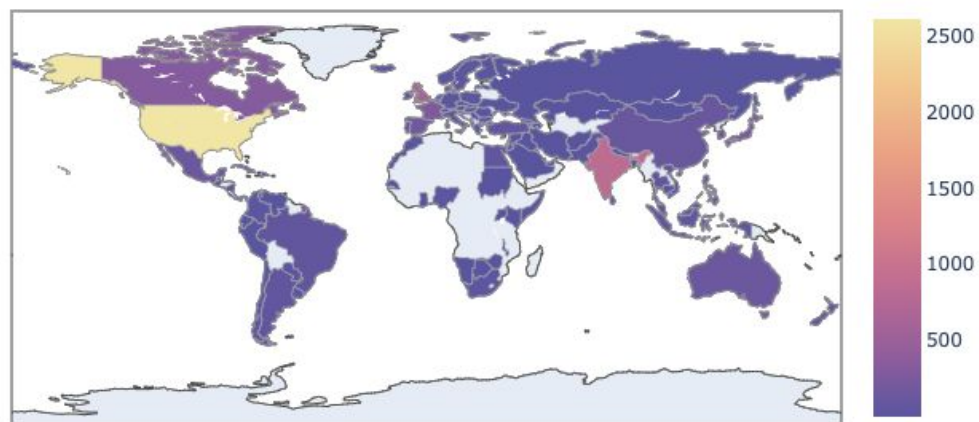
The United States looks to be the trailblazer in content production with over 2600 total content titles, followed closely by India with 838 total titles, and the United Kingdom with 601 total titles.

3. Who are the top Movie and TV Show directors on Netflix, how many are women?

Top 10 Movie Directors

Top 10 TV Show Directors
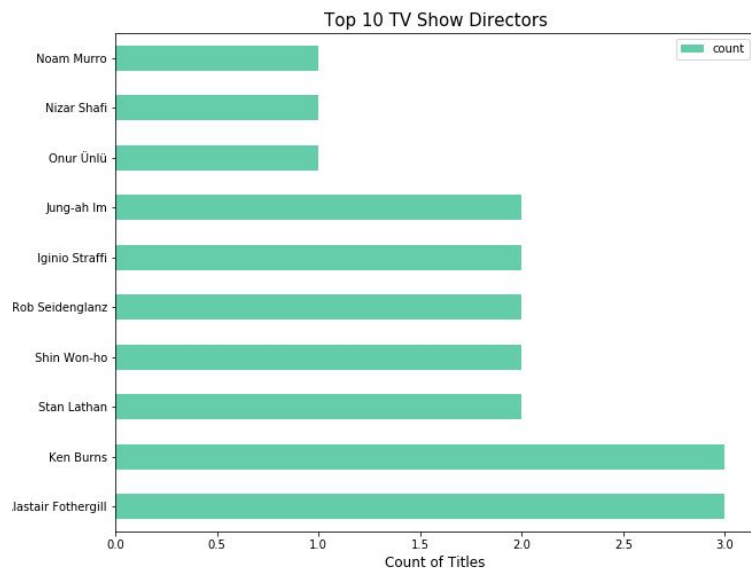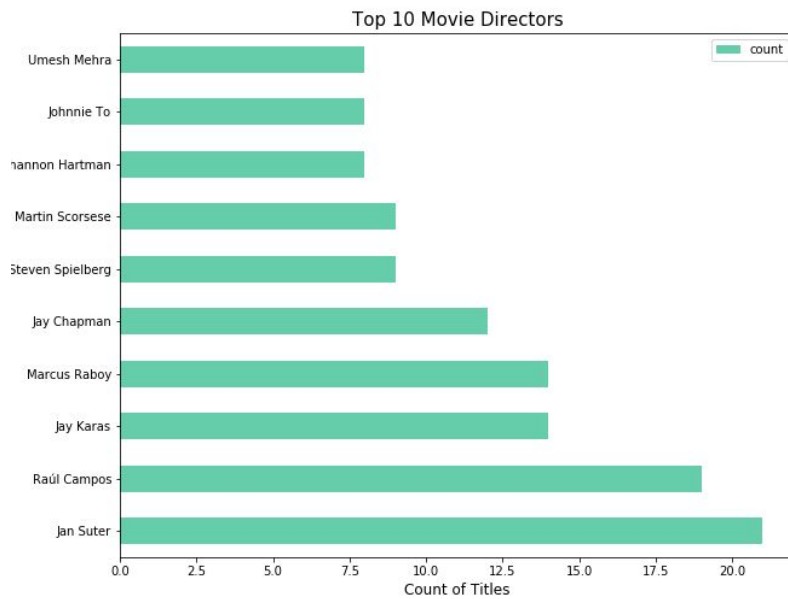
From the data it can be noted that multiple directors have collaborated for a single movie or a TV show. Of those, Jan Suter seems to have directed the most number of movies with 21 titles followed by Raúl Campos with 19 titles.

For TV shows, Alastair Fothergill and Ken Burns have directed the most number of TV shows with 3 titles each.

Directors:Other vs Female based on Type
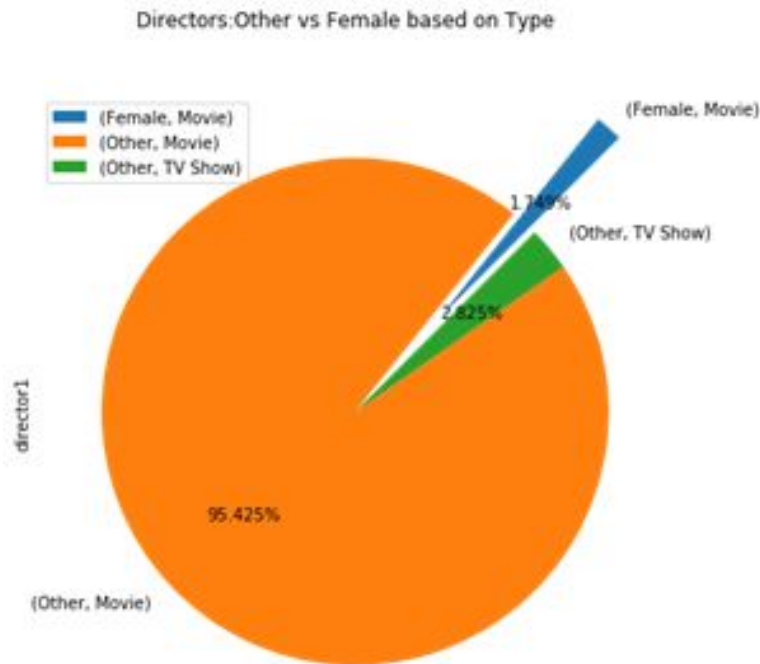
The above plot displays the ratio of Female Directors as opposed to Other Category (which includes Male and multiple directors) who have directed movies and TV shows. Female Directors only comprise 1.749% of the total directors population directing movies and TV shows. We see that Female Directors haven't directed any TV shows solely; they may have collaborated with other directors. The below plot excludes the directors who directed TV shows and we can see that the ratio slightly increases to 1.800%
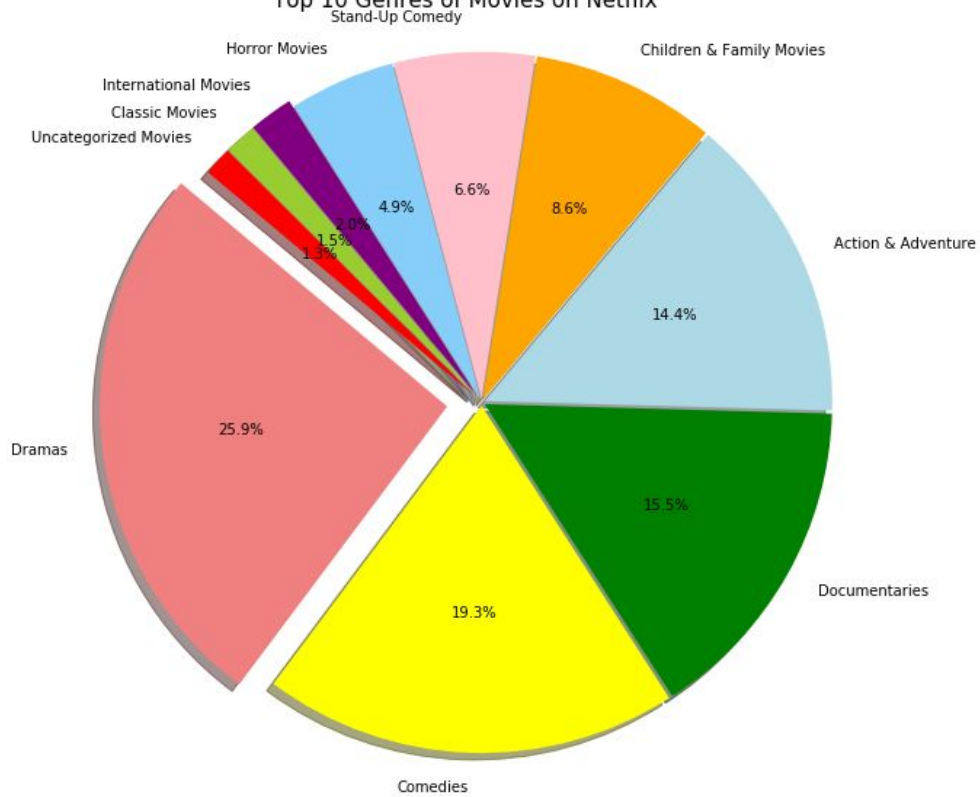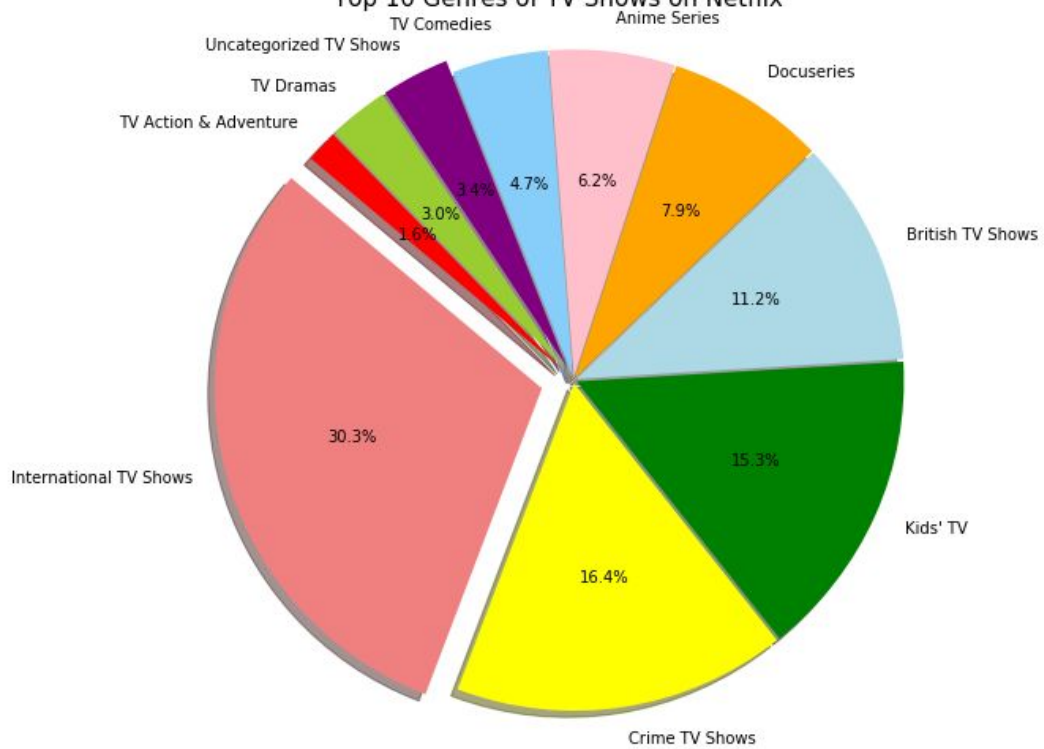
Genres/Ratings:

4. How is the content categorized?
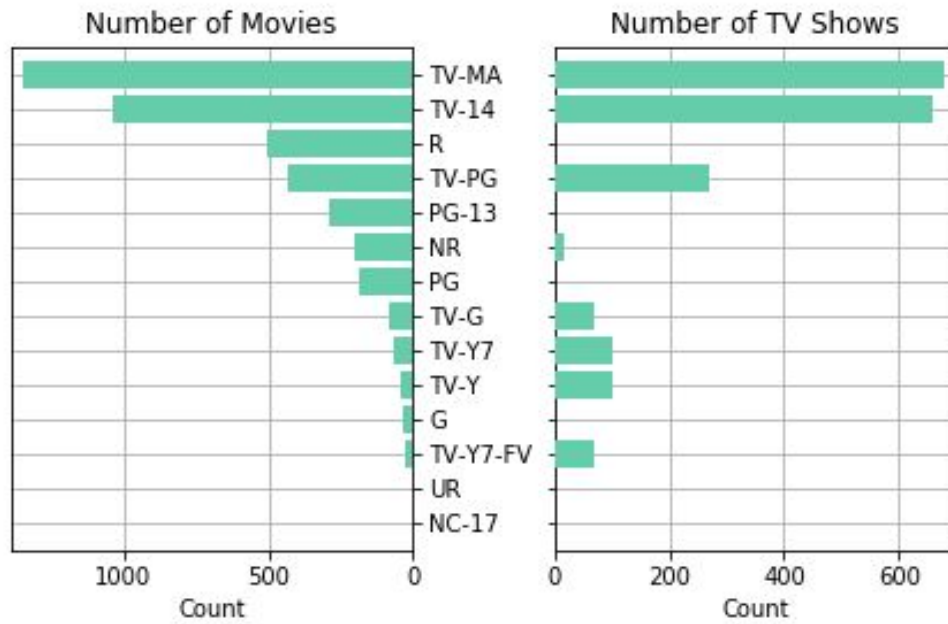
Most content offered within Netflix is Rated TV-MA, TV-14, R, TV-PG, and PG-13.

Top 10 Genres of Movies on Netflix

Stand-Up Comedy

Horror Movies

International Movies

Classic Movies

Uncategorized Movies

Children & Family Movies

4.9%

6.6%

8.6%

Action & Adventure

14.4%

2.0%

1.5%

1.3%

Dramas

25.9%

Documentaries

15.5%

Comedies

19.3%

Top 10 Genres of TV Shows on Netflix

TV Comedies

Anime Series

Uncategorized TV Shows

Docuseries

TV Dramas

TV Action & Adventure

4.7%

6.2%

7.9%

3.4%

3.0%

1.6%

British TV Shows

11.2%

International TV Shows

30.3%

Kids' TV

15.3%

Crime TV Shows

16.4%

For movies, Dramas (29%) have the highest percentage of content, followed by Comedies (21.3%) and Action & Adventure (16.2%). Of the top 10, the least amount of content are Thrillers (1.1%), Classic (1.5%) and International Movies (2.1%).

For TV shows, International TV shows (25.8%) have the highest percentage of content offered, followed by Crime TV Shows (20.4%) and British TV shows (18.3%). Of the top 10, the least content available for TV shows are Comedies (2.2%) and Action & Adventure (2.2%)

5. What genre of Movies do Females' direct?

Top Genre Categories for Female Directors
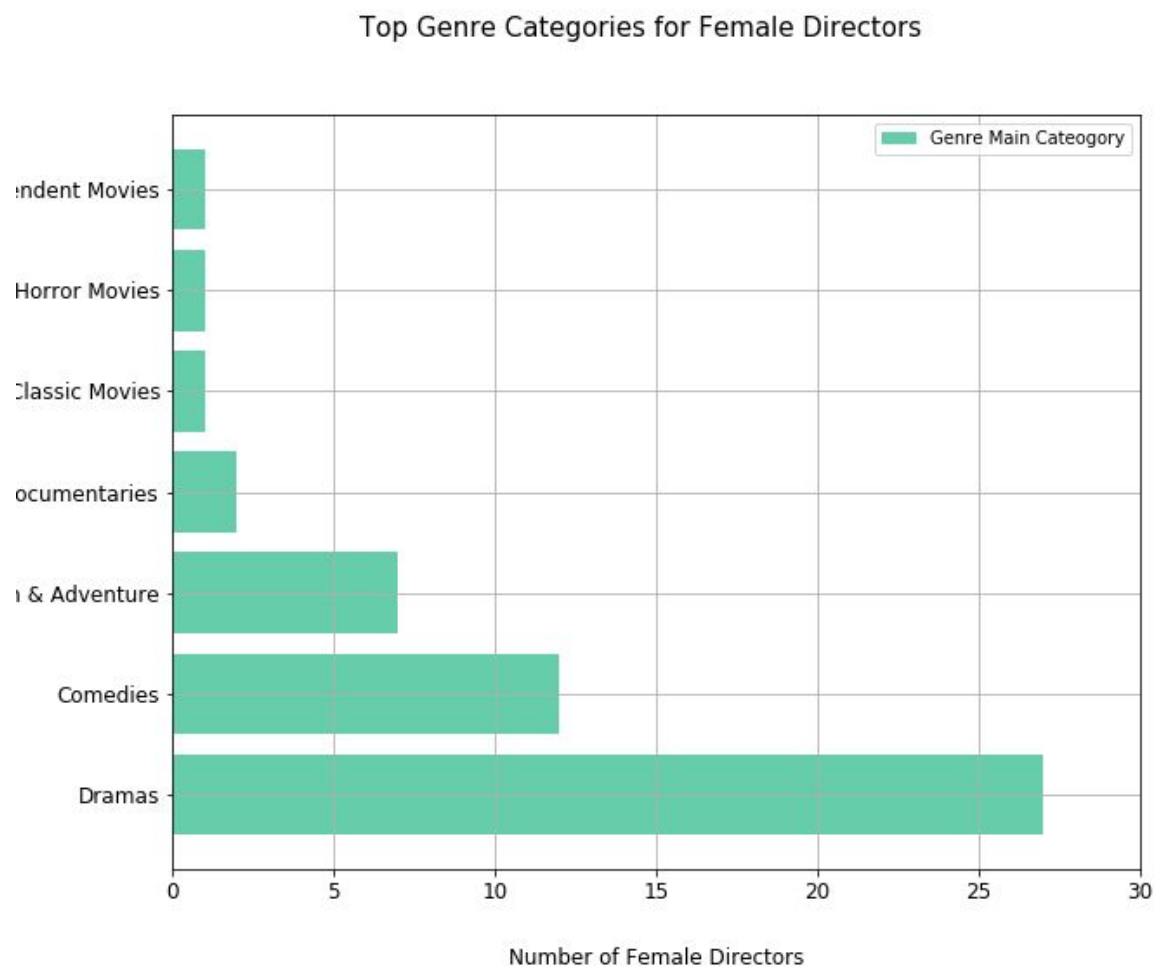
The above plot analyzes the relationship between the number of Female Directors and the Genre Category they directed. We see that 'Dramas' is the most likeable genre category for female directors or female directors are most likely to direct movies that are dramatic.

Duration:

6. Has the time of Movie Duration reduced over the period of time?

Movie Duration by Years
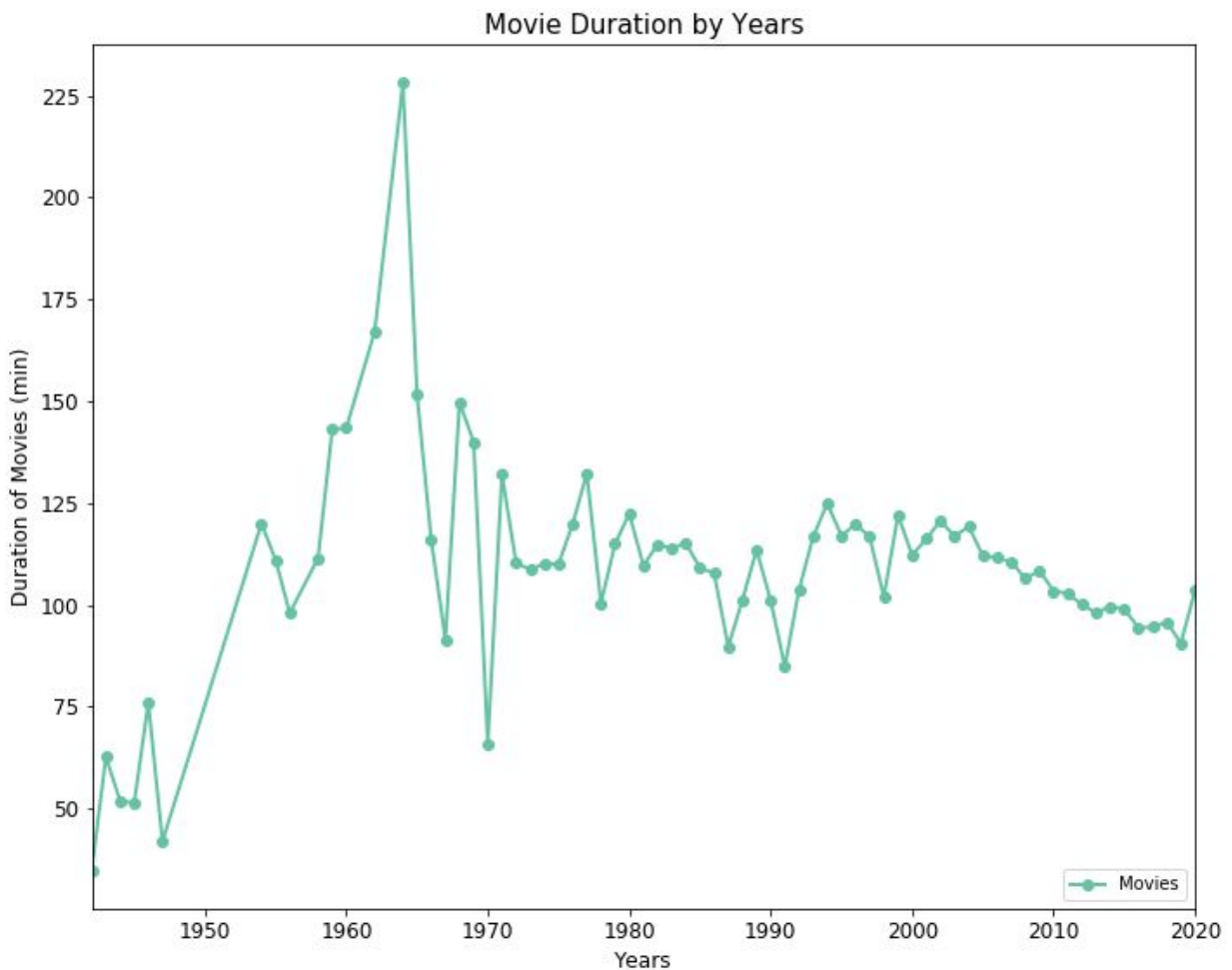
There is a substantial increase in the durations (feature length of films) from the 1940s up to mid 1960s. Duration decreases until the 1970s, then it remains relatively stable through 2020.

Duration of Movie Genres
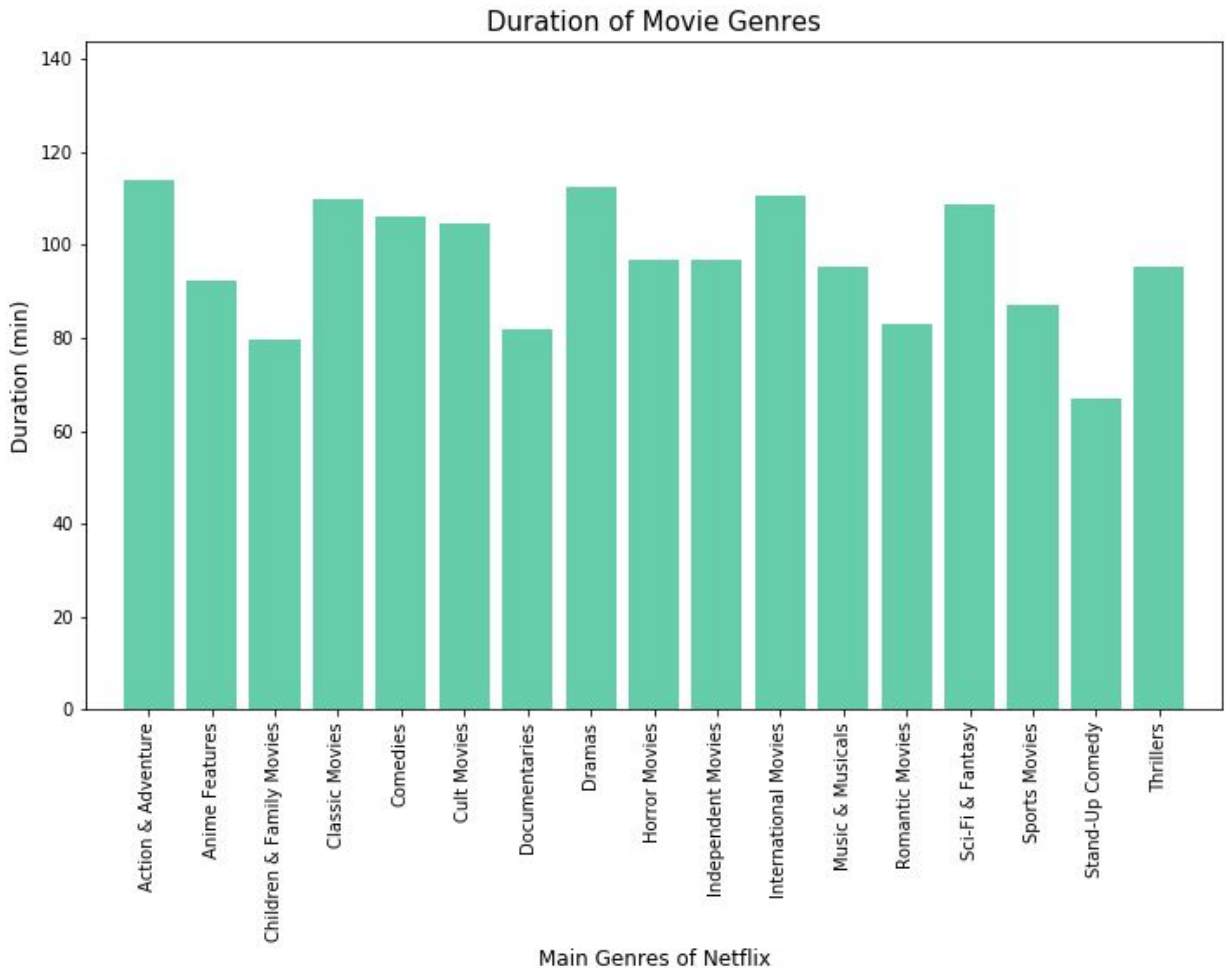
The plot above shows the average durations of Movie Genres on Netflix. Classic movies, Dramas, and International movies have the longest durations whereas Children & Family movies, Documentaries and Stand-up Comedies have the shortest durations.

There seems to be much variation in the length of the movies until the 1980s since these movies were generally Art and Cult movies. From the beginning of the 2000s, average duration has decreased approximately by 30%, probably because of internet usage and streaming platforms like Youtube.

Nearly half of all the movies belong to the genres of Dramas and Comedies. whereas Documentaries and Action movies nearly have the same share of, around 15%. This is an interesting outcome that shows people have preference for these two genres.

Duration of Movie Vs Rating
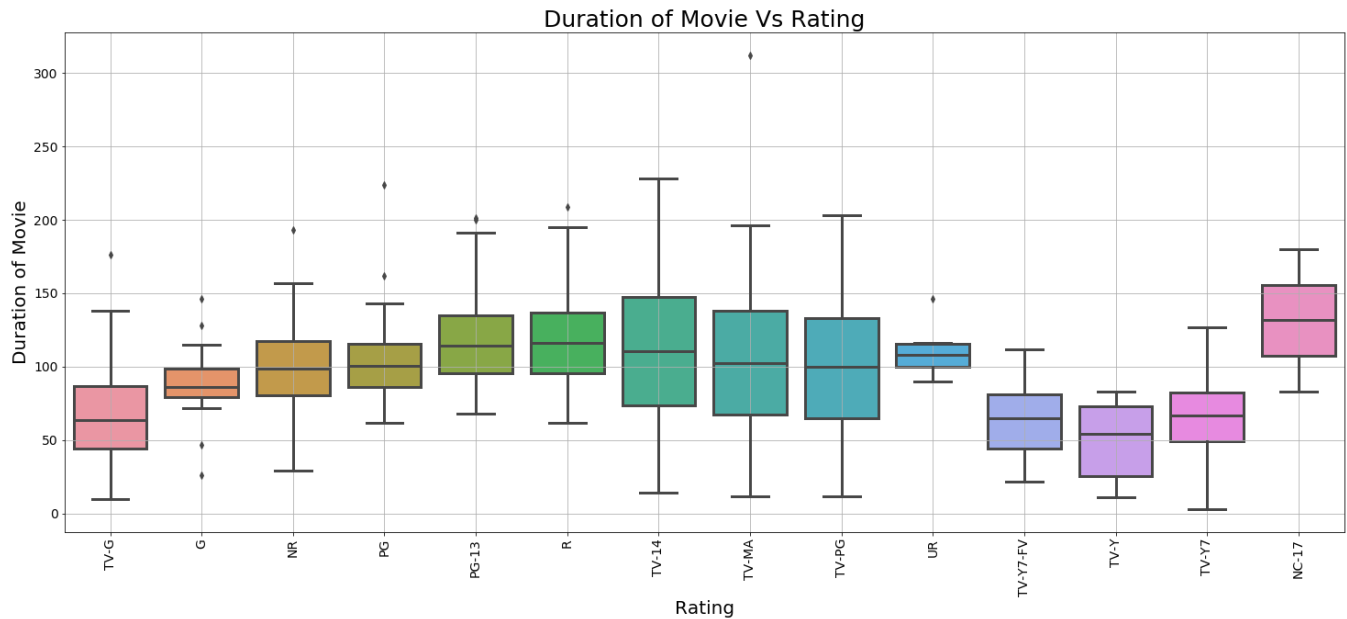
It is interesting to note that of all the movies added to Netflix, movies with the ratings TV-MA and TV-14 seem to be of the longest duration as well possibly suggesting age of the audience and duration of the movies might be related. However, more data would be needed to support such an observation.

The length of TV shows is expressed in terms of the number of Seasons, and no information is available regarding the length of a season in minutes. As such there is not enough information to chart a relationship between duration of a TV show and rating, similar to movies (as the above plot).

7. What is the most common duration of Movies Females direct?

Length of Movies directed by Female Directors
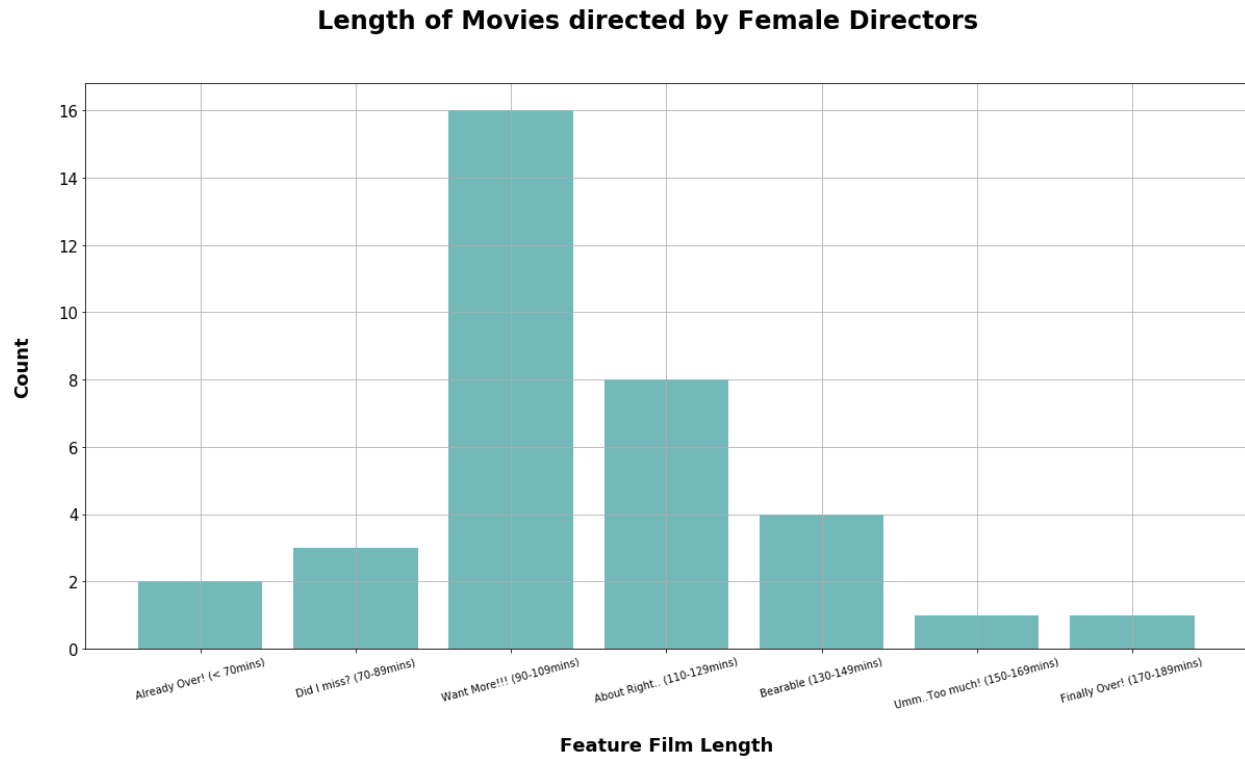
The above plot is analyzing the count of movies with durations directed by Female Directors. We see that Female Directors have directed more movies in the "Want More" category which is of duration from 70 to 90 min. They do not prefer to direct longer duration movies.

Additional Fun Findings:

Top 10 Movie Actors

Top 10 TV Show Actors
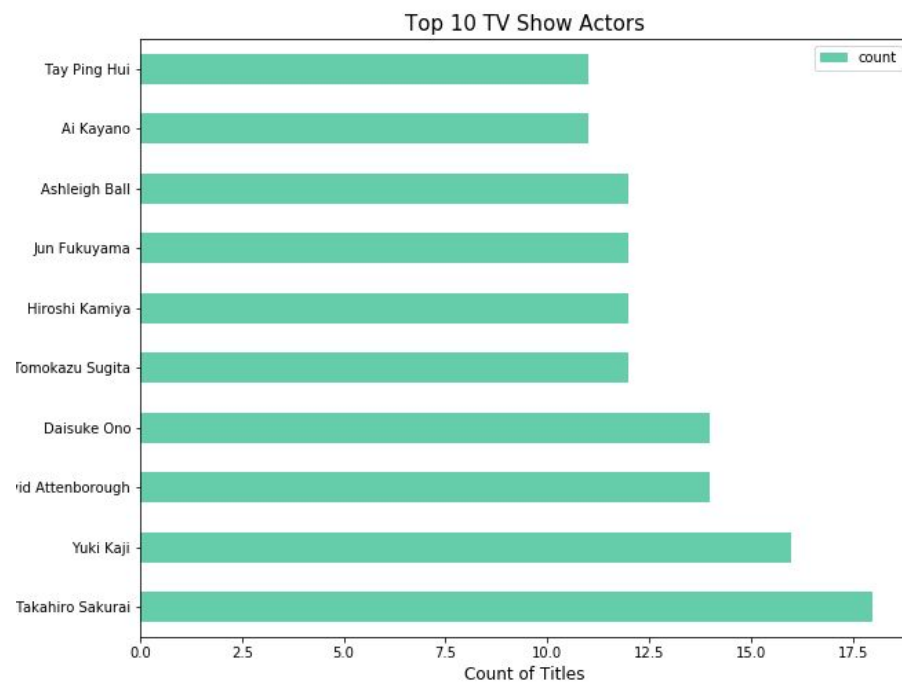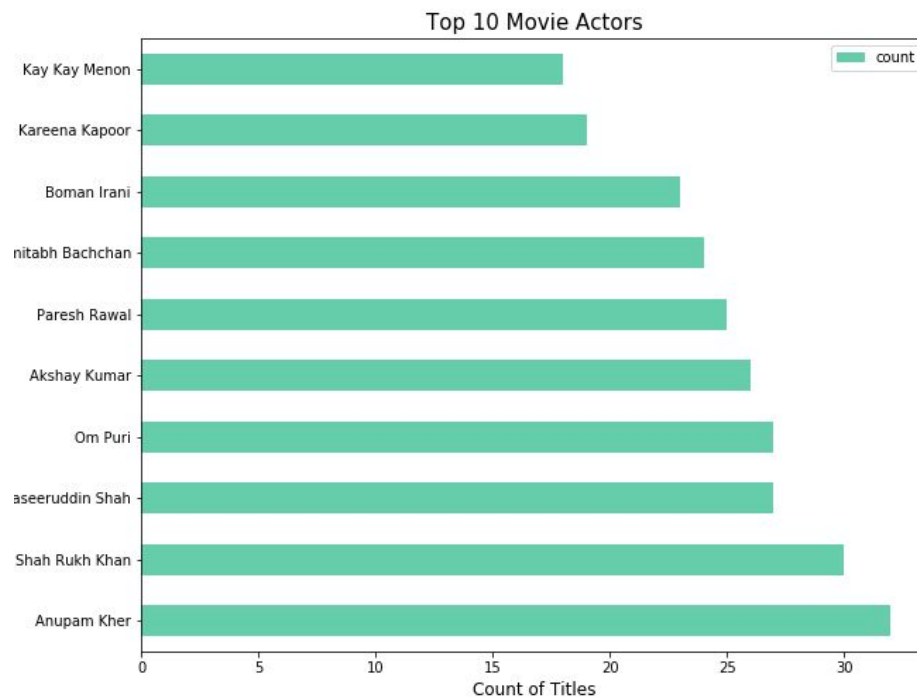
Anupam Kher seems to be the actor with the most number of movie titles, 32, followed by Shahrukh Khan with 30 movie titles. It is also interesting to note that all the actors listed above, (with the most number of movies) are from India.

Similarly, actor Takahiro Sakurai acted in the most number of TV shows with 18 titles followed closely by actor Yuki Kaji with 16 titles.

Most common words for Title
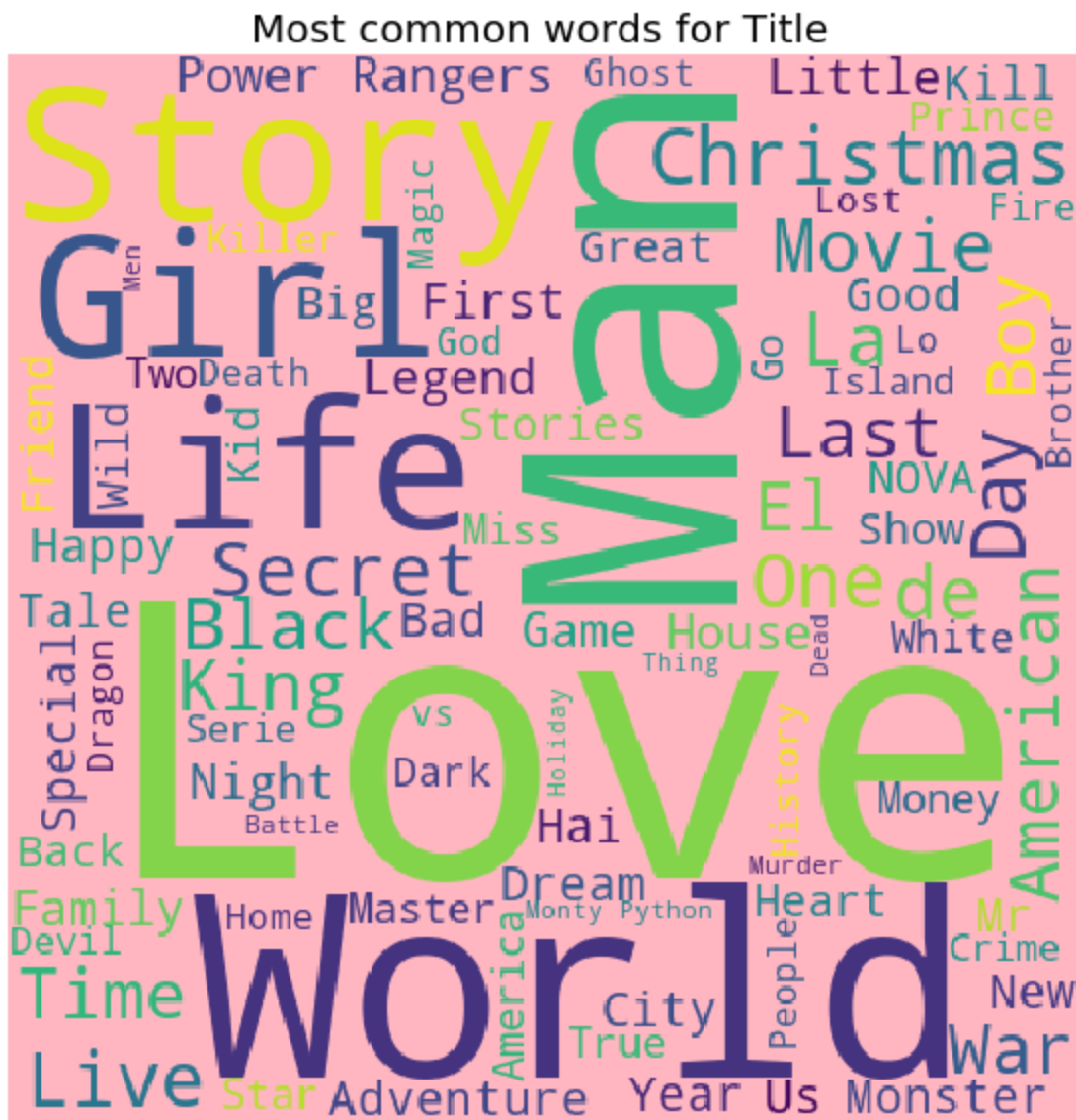
Love, World, Man are some of the most prominent words used for titles of the content added to Netflix.

Conclusion and Insights

- Netflix continues to increase its content year over year. Based on the trend, it will be higher in 2020 compared with 2019. The percentage of content added from 2018 to 2019 is higher for TV shows opposed to movies. There may be a shift in viewers' interest between TV shows and movies, more data is needed to confirm this.

- The genre 'Dramas' of movies has the most content and is also most preferred by Female Directors. It is also amongst the top 3 longest duration of movies.

- The majority of the content is for mature audiences (TV-MA) or audiences with ages 14 and above (TV-14).

- Female directors represent only 1.7% of the total number of directors.

- International TV being the highest count of TV Show offerings amongst all genre types.

- The number of movies and TV shows added by the United Kingdom are approximately the same as opposed to the other top 2 content producers US and India.