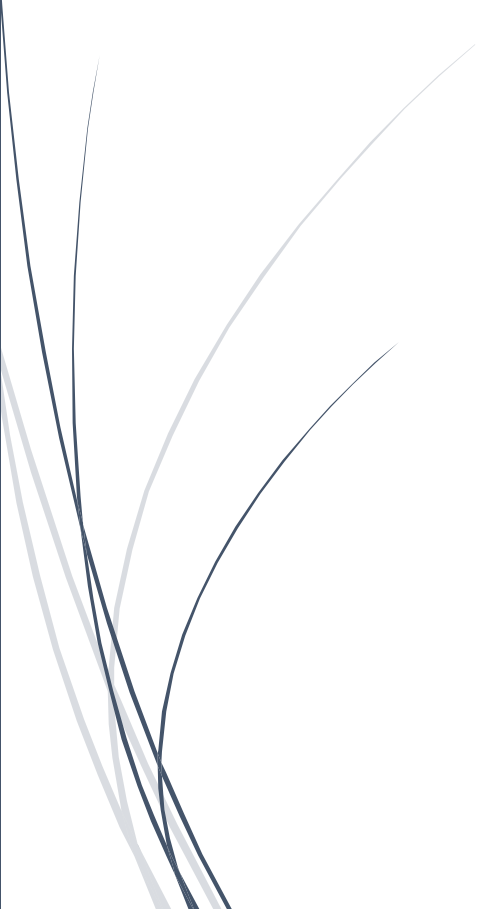


A dark blue vertical bar runs down the left side of the page. A blue arrow points to the right from the bar, containing the date.

04/25/2023

# Netflix TV Series Analysis

Several thin, curved lines in shades of blue and grey sweep upwards from the bottom left corner of the page.

Bradley Beck  
MDA 720

## Table of Contents

<b>Background .....</b>	<b>2</b>
<b>Objective .....</b>	<b>2</b>
<b>Data Manipulation .....</b>	<b>3</b>
<b>Sentiment Analysis .....</b>	<b>3</b>
<b>Data Visualization .....</b>	<b>4</b>
<b>Conclusion .....</b>	<b>8</b>

## Background

Becks Entertainment is prepared to launch its first television series on Netflix in the year 2025. As a new company in the entertainment business, it is necessary that our analytics department configures the best research on what has or has not allowed other series to gain success on Netflix. With the information our analytics team gathers we will have better insights on what Becks Entertainment's new show should consist of. Our team of analysts have manipulated a [Netflix](#) dataset that was obtained from a website known as Kaggle. This dataset consists of 12 columns and 8,793 rows. The 12 columns in the dataset are show ID, type, title, director, cast, country, date added, release year, rating, duration, listed in, and description. With the information offered our team will provide a thorough analysis on tv shows released up until 2021. The data contains information for both tv shows and movies although our team is only focusing on shows as of now. If Becks Entertainment decides to create a movie in the future our analyst would need to reevaluate this data.

## Objective

Solely in the year 2021, over 300 tv series were released on Netflix. The goal of Becks Entertainment is to create a show that separates itself from the rest. This will be achieved by creating a sentiment analysis as well as multiple visualization to determine what majority of people are looking for in a television show. The sentiment analysis will analyze the description column of the dataset to separate positive, negative, and neutral tv show descriptions. This will offer Becks Entertainment a better understanding of how to word show descriptions for the best result. Next, our team will create various visualizations to obtain data on the most popular genres, ratings, and directors being viewed on Netflix. Following this a word cloud will be

constructed from the title column to determine the most popular words used in show titles. This is important so we can give our show an eye-catching title. With all this data put together, Becks Entertainment will have a better idea of what is necessary to produce a popular show.

## Data Manipulation

Prior to running the sentiment analysis or any visualizations, it is important to manipulate the dataset. The first step in doing this was to check for any non-applicable data in each of the columns. Our findings were that the director, cast, country, rating, date added, and duration columns all contained rows with non-applicable data. The solution to this was to drop all non-applicable data found in the rating and date added columns. However, in the remaining columns our team chose to replace the non-applicable data rather than simply removing it. This was done by putting “No Director” in place of all NAs in the director column, “No Cast” in the cast column, “Country Unavailable” in the country column, and “No Duration” in the duration column. After this was complete all the non-applicable data was successfully removed or filled in. Next, our analysts retrieved more information on the dataset, such as the data types of each column. Everything checked out nicely and it was now time to dive deeper into the data.

## Sentiment Analysis

After properly manipulating the dataset, our team decided to perform a sentiment analysis on the description column. This technique is used to extract subjective information from a categorical column and categorize it into positive, negative, or neutral statements. Performing this analysis allows Becks Entertainment to gauge a better understanding on the public opinion

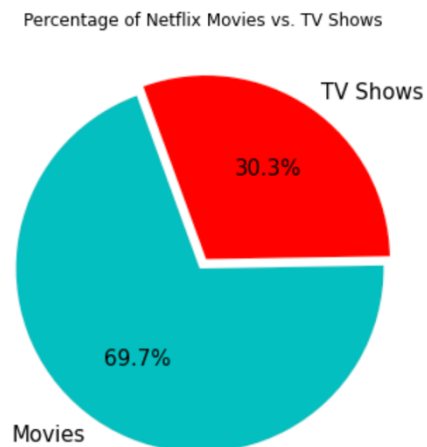
of previous television show descriptions. From this data our analysts can review all the positive, negative, and neutral descriptions given and determine what contributes most notably toward positive results. This will ultimately allow the company to generate the best possible description for our own television show to ensure it falls in the positive category. Developing a show description that can land in the positive is very important to attract the most attention. The results of the analysis can be observed below.

	release_year	description	Sentiment
1	2021	After crossing paths at a party, a Cape Town t...	Neutral
2	2021	To protect his family from a powerful drug lor...	Negative
3	2021	Feuds, flirtations and toilet talk go down amo...	Negative
4	2021	In a city of coaching centers known to train I...	Neutral
5	2021	The arrival of a charismatic young priest brin...	Neutral
8	2021	A talented batch of amateur bakers face off in...	Positive
10	2021	Sicily boasts a bold "Anti-Mafia" coalition. B...	Positive

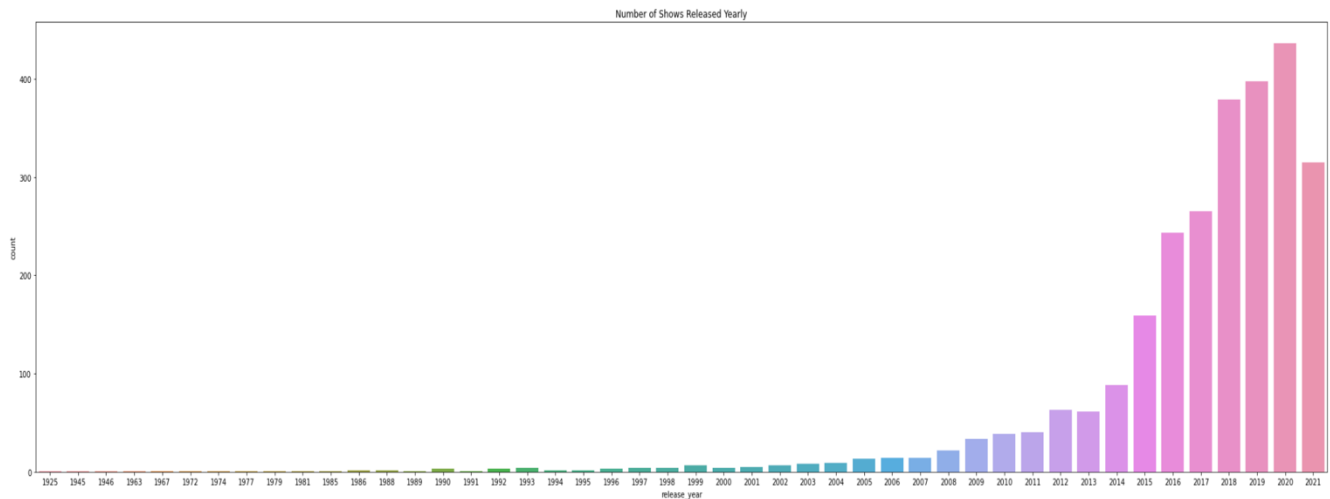
## Data Visualization

Netflix is a major platform that is used across the world for individuals to watch their favorite tv shows and movies. As of 2021, Netflix contains a much larger movie selection compared to television shows. The difference in number of movies versus television shows is demonstrated below. As observed in the pie chart,

Netflix contains 69.7% movies and only 30.3% television shows. Our company is expecting to join the slim percentage of television shows in the year 2025. Now that we know the rate of shows that can be accessed on Netflix, our team decided to investigate the number of shows that are



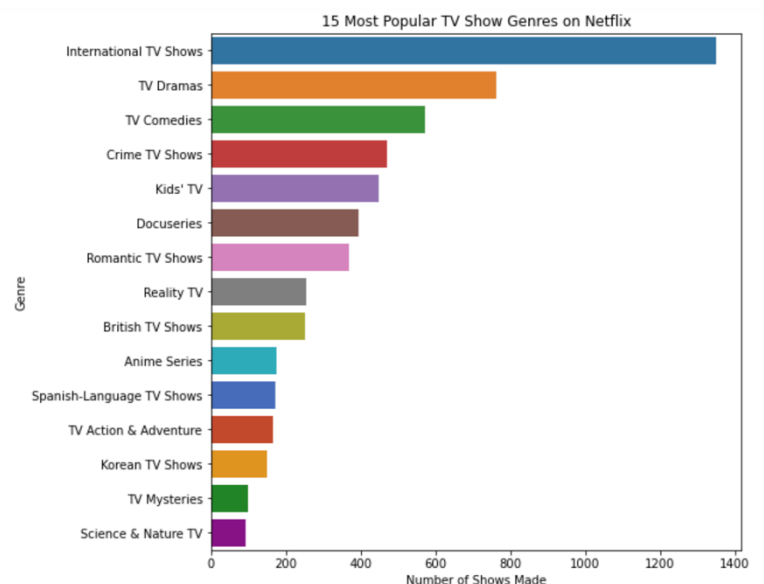
released each year. Knowing this number will give us an idea of how many competitors to expect.



Up until 2020 the number of shows produced each year continued to increase, although in 2021 this number dropped off a bit from the previous three years. Even though it only decreased a small amount, it is still much greater than any year leading up to 2017. With the magnitude of shows being released annually, it is important for Becks Entertainment to further our analysis so we can stand out from the rest. In doing

so our team decided to figure out the most common television show genres on the platform. These findings can be analyzed in the bar chart shown below.

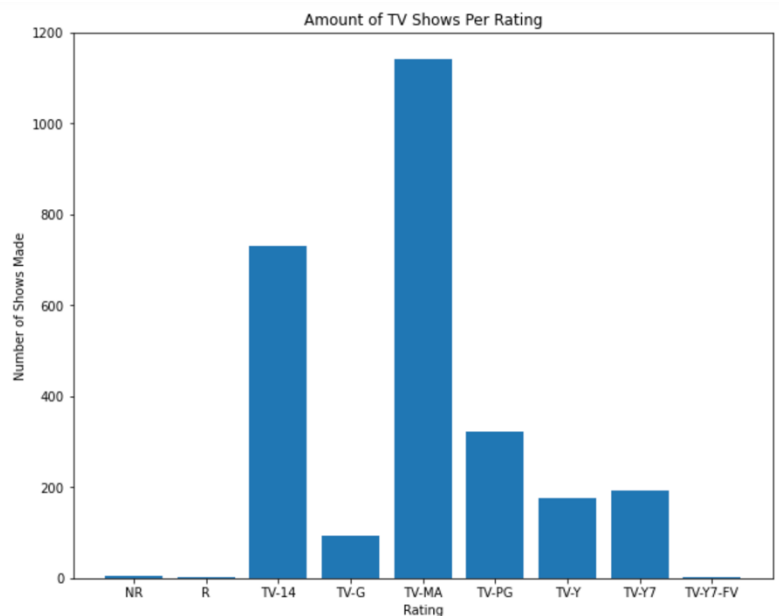
Behold, the five most frequent television genres are international shows, dramas, comedies, crimes, and kid shows. While



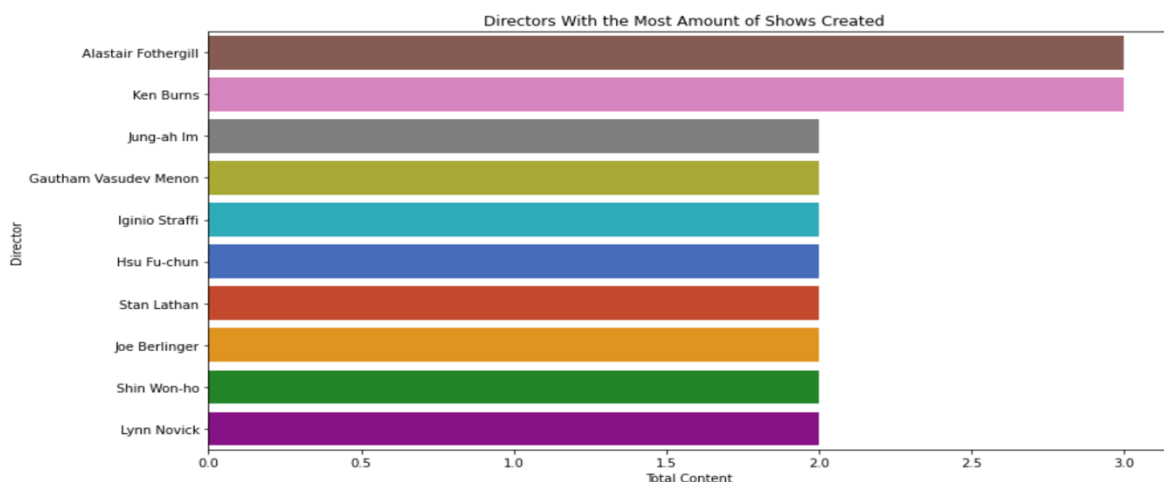
these may not be the most popular genres watched by viewers, they are the most produced genres. With this information we can assume that releasing a show in one of the top five genres will result in a lot of competition.

Following these findings, our team investigated what rating majority of shows fall into. We can utilize this information to target an audience that may not have many show options based

on the content they desire. The better part of Netflix shows fall under the rating TV-MA. These are shows that are intended to be watched by viewers 18 and older. Based off this data a rating such as TV-G would be a great audience to enter. With less than 100 shows in this

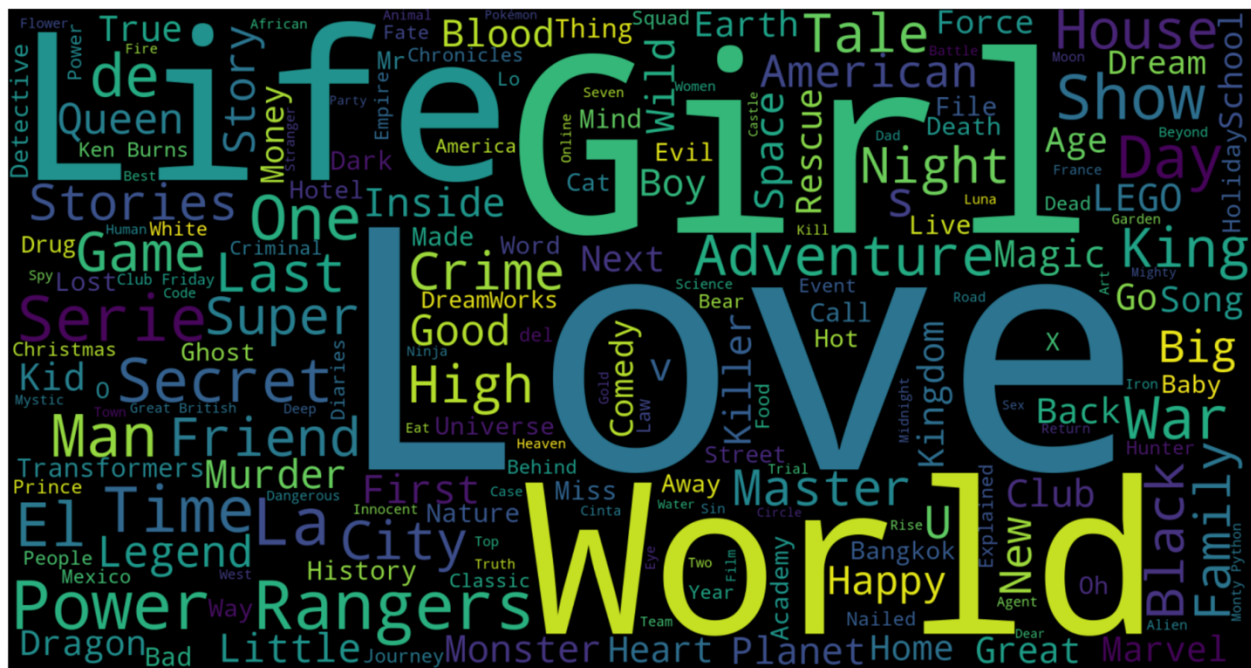


rating we would be putting ourselves in a category with limited competition. On top of this, shows that are categorized in this rating are meant for all ages and gives Becks Entertainment the ability to market a family friendly television series for all to enjoy. With this being the first



television show created by Becks Entertainment, our company feels it is important to hire a well-known director with experience. To pinpoint this information our analyst put together a bar graph that demonstrates which directors have produced the most shows.

With the information obtained we will be reaching out to Alastair Fothergill and Ken Burns in hopes that they are available to work with us. If neither of them is available, we will move on to one of the other experienced directors seen above. A director with knowledge of the television world would be extremely helpful towards reaching our goals. Lastly, in our analyst exploration they have put together a word cloud from the title column. This word cloud will be used to narrow down some of the most popular words seen in Netflix tv show titles.



The larger the font indicates that the word is used more often in show titles. For example, words such as life, girl, love, and world appear much more frequently than words like show, friend, and crime. Using this knowledge Becks Entertainment will gain inspiration towards creating a unique title that will draw in viewers' attention.



## Conclusion

With the research conducted by our team of analysts, Becks Entertainment can now develop a base plan on what our show will entail. The development of our show will be categorized in a genre outside of the top five to increase our chances of success. By choosing a less common genre we are separating ourselves from more established competitors within. Other than this we will make the series fall under the TV-G rating. With such a limited number of shows in this rating we will be able to attract a very exclusive audience. Our findings have also allowed us to narrow down a list of directors. Targeting a well-established director will allow us to reach better actors and provide us with instant attention. Once all this information is established, we will use the data from the word cloud to create a unique and stand out name for our series. Finally, the data we retrieved from the sentiment analysis will give us the best opportunity to create a positive show description. This is very important as most people read a shows description before deciding whether to watch. From the data retrieved in this analysis, Becks Entertainment has gained valuable insight on what it will take to release a show that attracts viewers.

## Works Cited

<https://www.kaggle.com/datasets/mahmoudtaya/netflixtitles>

<https://github.com/bbeck15/MDA-720-Applied-Business-Analytics-Capstone-Project/blob/main/MDA720%20Capstone.ipynb>