

## Problem Statement:

Descriptive analysis of YouTube trending videos in US based on their Title, Likes, Dislikes, Comments, Tags, Publish Time and Description.

## Methodology:

1. Collection of datasets from the Kaggle.
2. Exploring the dataset to find out what are the variables we are having in our dataset.
3. Try to find out the articles about the YouTube trending video on the internet like which attributes (does number of views makes a video Trending or a video from a familiar channel causes trending) makes a video trending.
4. Cleaning and organizing the different datasets into one final dataset which will be used for our analysis.
5. Take decision on which type visualizations to be performed based on the data so that we could show and explain to the listeners.
6. Python 3 environment was used to plot the graph containing likes and the time durations which the videos were posted.
7. Few libraries such as numpy, pandas, scipy, seaborn, and matplotlib were used to make the visualizations and even Tableau.

## Data Cleaning:

Most of our dataset is clean. But we have to arrange the different datasets like joining based on the variables in common. We had first explored each CSV and JSON files just by opening them directly rather than exploring through python because excel is also a powerful tool for handling the missing data or the applying basic stat functions like mean, average, sum etc. Trending dates were separated into date, time, and hour. Published time was divided into month, day, and day of the week.

## Research about the Project:

The main purpose of this research was to find out what aspect makes the Channels more trending in YouTube. YouTube uses some kind of algorithm (black box) to decide which video to add in the trending list. Like it considers the length of the video and even the session of the user which plays an important role like how much time did the user spent on watching the video. Number of views the video had reached with in the certain period of time. Most of the times they even try to avoid the misleading videos, but it also depends on the viewers and they have to report. Until and unless they are not reporting the video cannot be removed from the site which ultimately can become the trending video. To overcome these types of misleading videos it said that it hired 10,000 people for a moderation team to try and combat this type of content, but even then, it's unclear how it will be deployed.

## Results and Findings:

The research showed what time durations are the best to publish videos that will have a higher probability to be in the trending list.

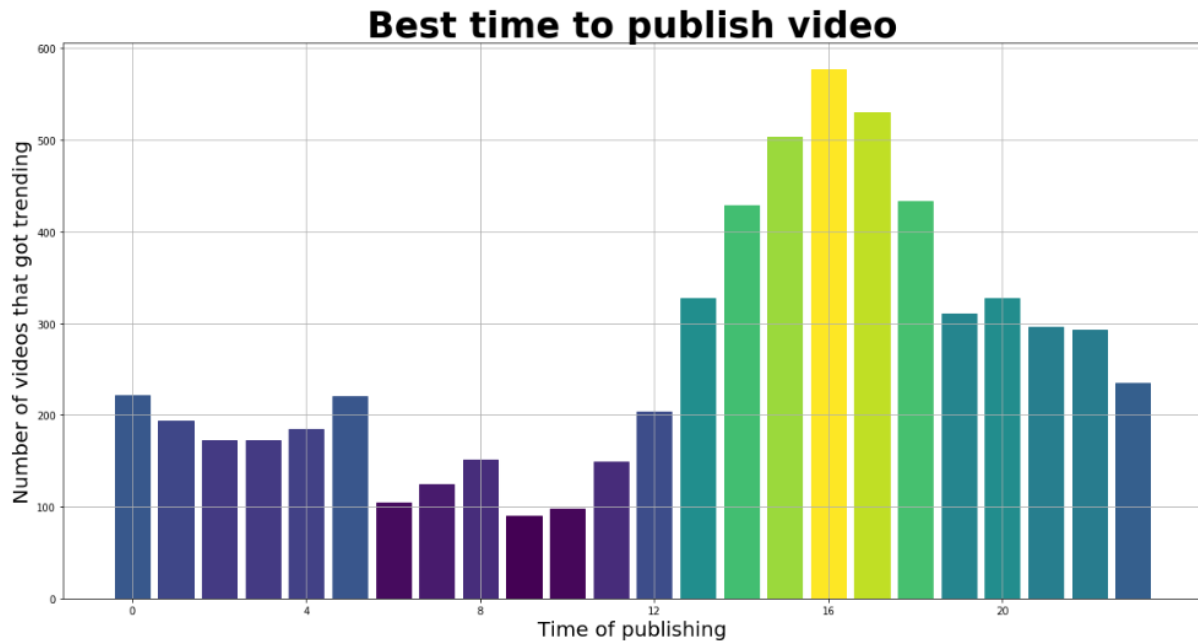


Fig 1: Number of videos that get trending vs Time of Publishing

This shows that around 4 pm in the evening is the best time to publish videos which will gain the most likes and pushed into the trending list.

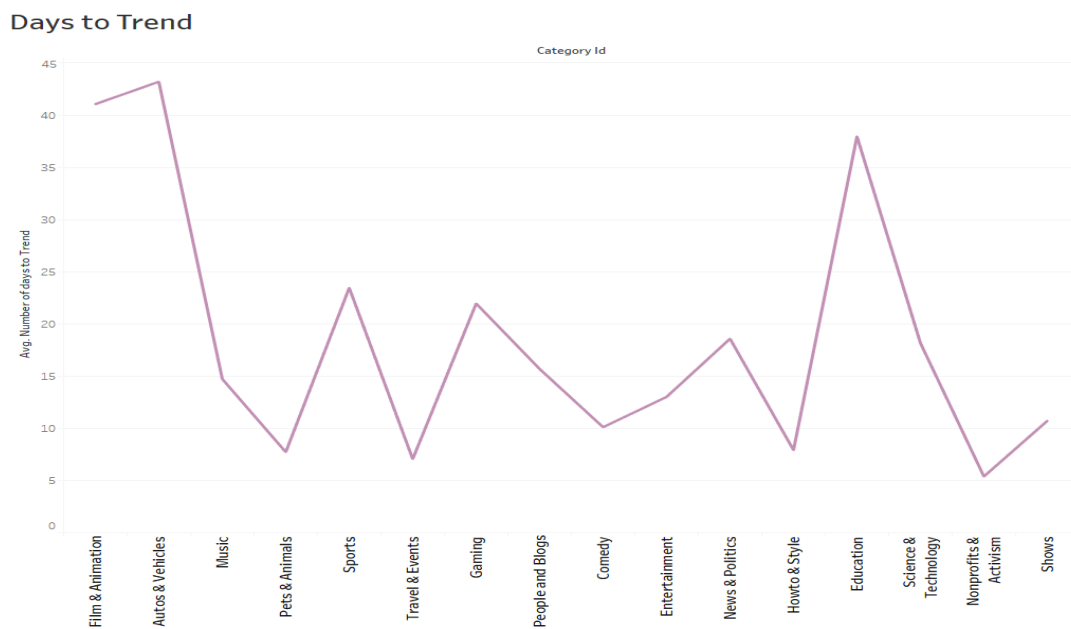


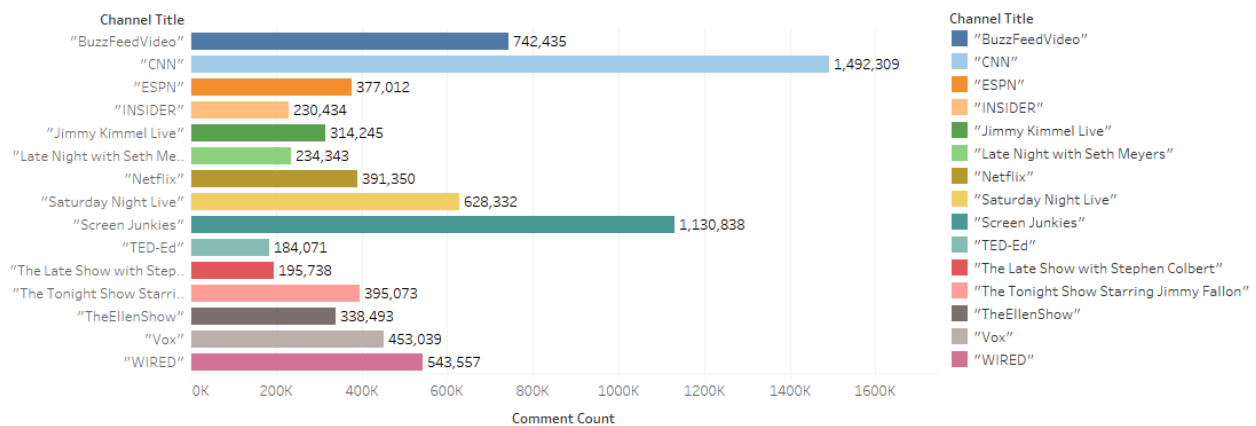
Fig 2: Channels vs Average Number of Days to Trend.

This graph shows the what type of videos takes least amount of days to trend in youtube. This graph shows the average days that a video takes to trend. It shows that Non-Profits and Activism takes least number of days to trend.

Travel & Events and Howto & Style takes the second place in the average trending days. Autos & Vehicles category takes maximum amount of days which makes auto and vehicles videos with least likes or views.

Comparison of Channel Title was done with three different variables, comments, likes and views. The result was, based on comments the news channels trends faster, based on likes movie review/entertainment channels trends faster, and based on views individual artists/entertainment channels trends faster.

Trending Channels Based on Comments in US

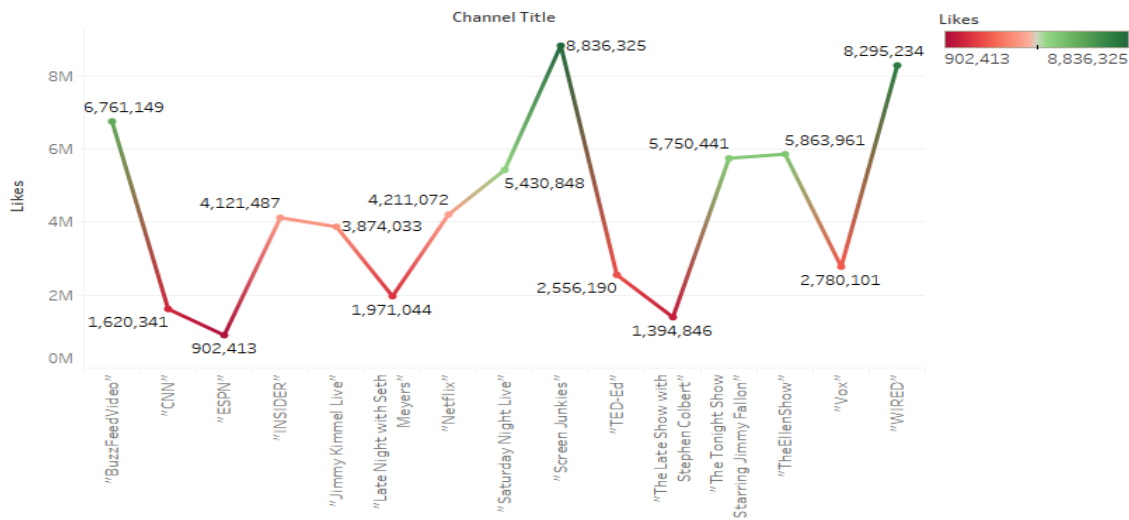


Sum of Comment Count for each Channel Title. Color shows details about Channel Title. The marks are labeled by sum of Comment Count. The data is filtered on Trending Based on Comments, which keeps 15 members.

Fig 3: Channel Title vs Comment Count

From the above visualization based on Total comments In the Channel we can see "CNN" is the most trending channel with around 1.5 million total comments.

Trending Channels Based on Number of Likes in US

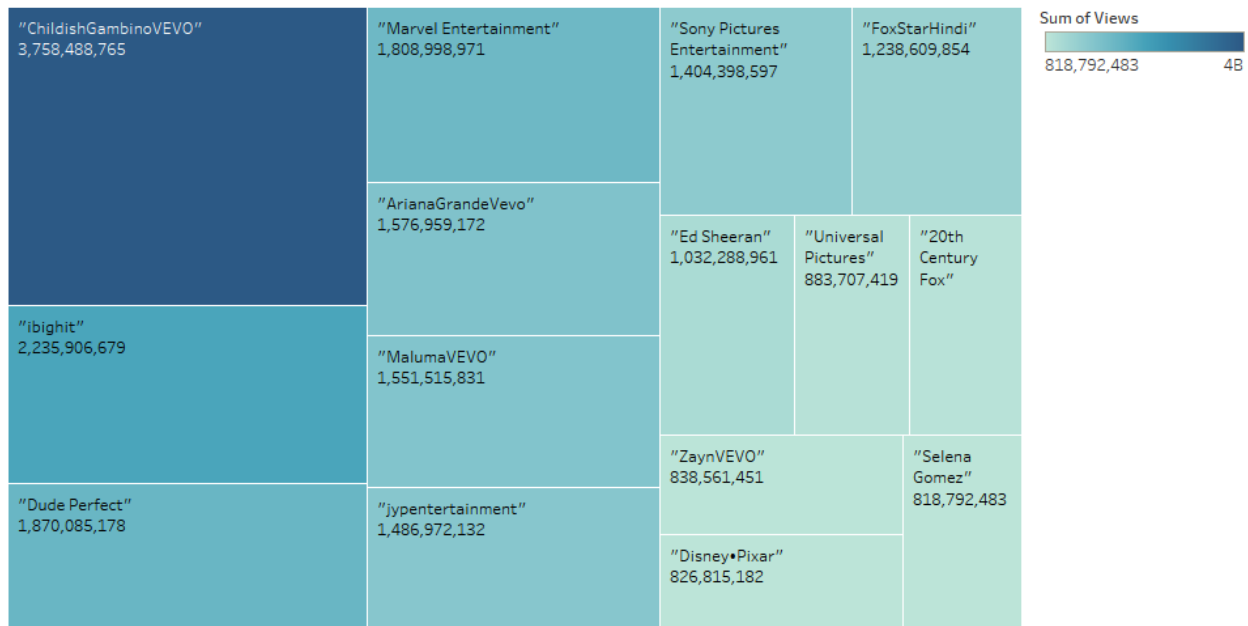


The trend of sum of Likes for Channel Title. Color shows sum of Likes. The marks are labeled by sum of Likes. The data is filtered on Trending Channels with most likes, which keeps 15 members.

Fig 4: Trending Channels based on Number of Likes

From the above visualization based on total likes we can see for the entertainment channels like **“Screen Junkies”** and **“WIRED”** the total likes are more than 8M.

Trending Channels Based on Number of Views in US



Channel Title and sum of Views. Color shows sum of Views. Size shows sum of Views. The marks are labeled by Channel Title and sum of Views. The data is filtered on Trending Channels with most views, which keeps 15 members.

Fig 5: Trending Channel Based on Number of Views

From the above visualization based on total views we can see for the individual artist channels like **“ChildishGambinoVEVO”** and **“ibighit”** the total views are more around 4 Billion and 2.3 Billion respectively followed by **“Marvel Entertainment”** with around 1.9 Billion views.

Sentimental analysis based on the Description about the video. This analysis is performed to know what kind of videos are mostly trending

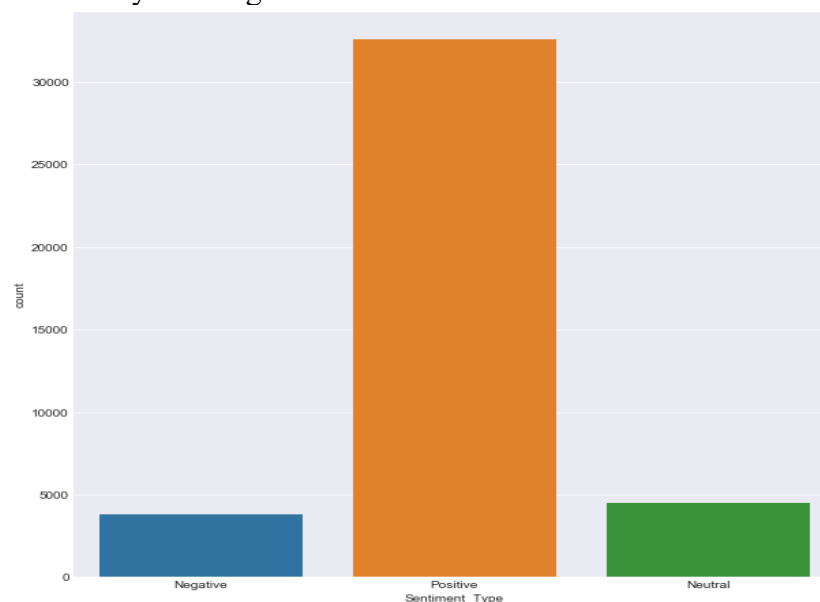


Fig 6: Sentimental Analysis Based on Description of video.

From the image we can see that videos with positive description have trended more in comparison with negative and neutral description. This states that most of the videos have the positive description.

Sentimental analysis based on the Tags of the videos. This analysis is performed to know which kind of tags are mostly trending.

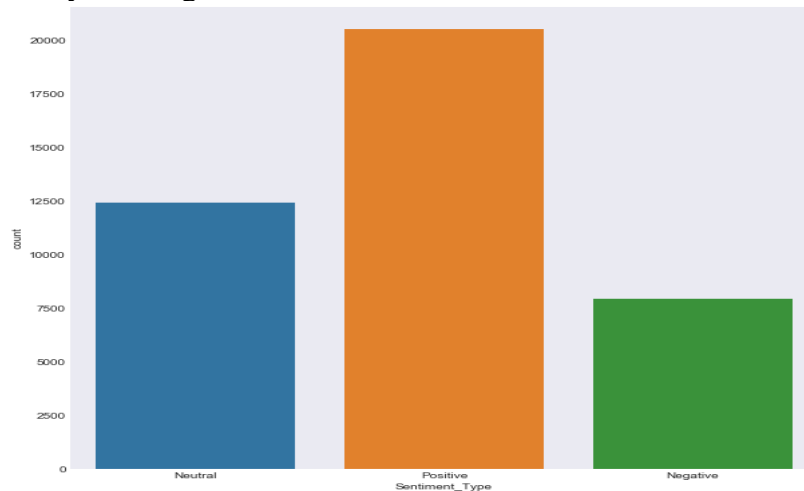


Fig 7: Sentimental Analysis Based on Tags of Videos.

From the image we can see that videos with positive tags have trended more in comparison with negative and neutral tags. But we can even see videos with neutral tags are also large in number. This states that most of the videos have the positive Tags and also neutral tags.

Sentimental analysis based on the Title of the videos. This analysis is performed to know which kind of Titles are mostly trending.

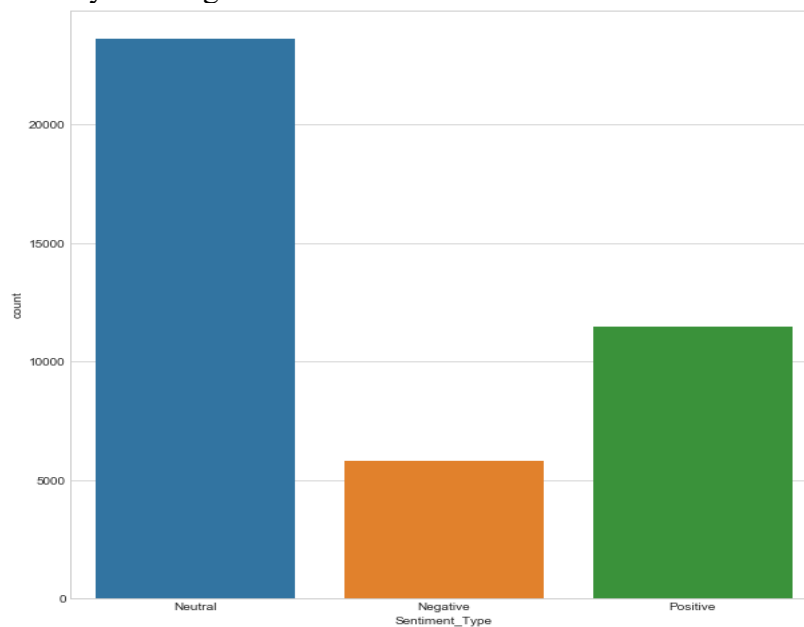


Fig 8: Sentimental Analysis Based on Title of the Videos.

From the image we can see that videos with neutral tags have trended more in comparison with negative and positive Titles. This states that most of the videos have the neutral Titles and then the positive Titles.

So from all these visualizations of sentimental analysis of Description, Tags and Titles. We can say that most of the videos with positive description have high trending videos but coming to Tags and Titles we could make decision that most of them are Neutral but a little tend towards positive sentiment.

#### Limitations:

1. We got dataset of list of trending videos. But they gave characteristics of each trending video. If they have provided some more numerical fields, we could have at least tried to find out algorithm.
2. We tried to work only for US region. If we could have worked on other regions, we could have got more interesting visualizations like how a channel performed in different regions.
3. Since we have dataset of all trending videos we couldn't get any data about non trending video and even there will be millions of videos which are being uploaded every day.
4. By seeing the results from our analysis it is clearly showing that our results are just apart to consider for trending videos but there are lot more factors which are considered in real.
5. The tags of the videos are filled up with unrelated and unwanted tags which can mess up the data cleaning and the analysis.
6. The dataset was missing the information like the number of times the channels shared in different social media, which could be used to determine if the channel trends based on sharing.

#### References:

1. <https://www.wetheunicorns.com/features/bad-baby-freak-creepy-trending-milk/>
2. <https://www.polygon.com/2018/2/21/17035758/how-does-youtube-trending-work-parkland>
3. <https://boostlikes.com/blog/2017/09/ways-video-trending-youtube>