

Analysis Results

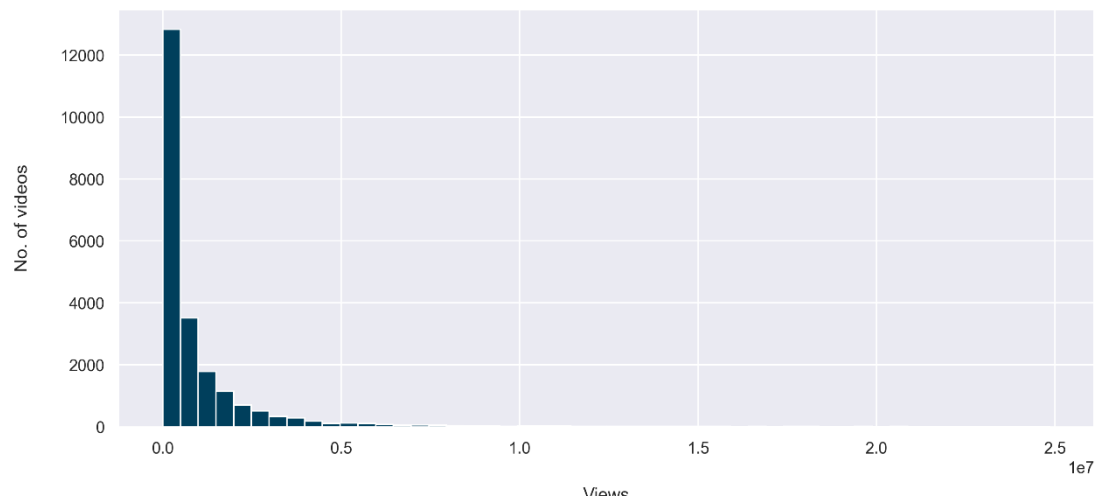
Summary of Some Key Results of the Analysis

- The dataset was collected in 2017 and 2018 with 61% of it in 2018 and 39% in 2017.
- The dataset contains collected data about trending videos over 125 days.
- The average number of views of a trending video is 1329987.
- The majority of trending videos have 40000 likes or less.
- ESPN was the channel that produced most trending videos.
- Words like “official”, “video”, “trailer”, “music”, “show”, “new”, “live”, “first”, “time”, and “HD” were among the most common in trending video titles.
- Almost all trending videos use tags. On average, 13 tags.
- Comments disabled videos had less views.
- 139 trending videos out of 22385 prevented users from rating
- There are just 80 trending videos that have both comments and ratings disabled.
- The most views came from the state of WA and city with large number of vies is Pensacola.

Trending Videos ,Views and likes

- The Trending Video with the Highest Number of Views is from Entertainment category with title “YouTube Rewind: The Shape of 2017” published on 2017-12-06 with highest number of views of 149376127 trending in December, 2017.
- The video with least number of vies is from News and Politics with title “1 dead, others injured after Ky. school shooting”.

And now we will look at the distribution of views for trending videos. The histogram below shows this distribution:

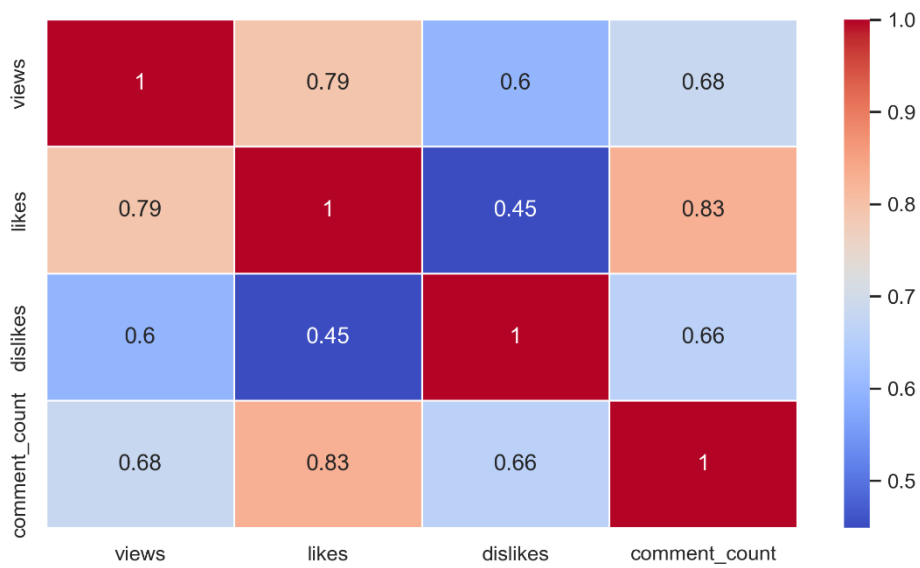


Now we see that the majority of trending videos have 1 million views or less. So, the percentage of videos have views 1million or less is 72.71%. Similarly, the percentage of videos with less than 5 million views is around 95.2%.

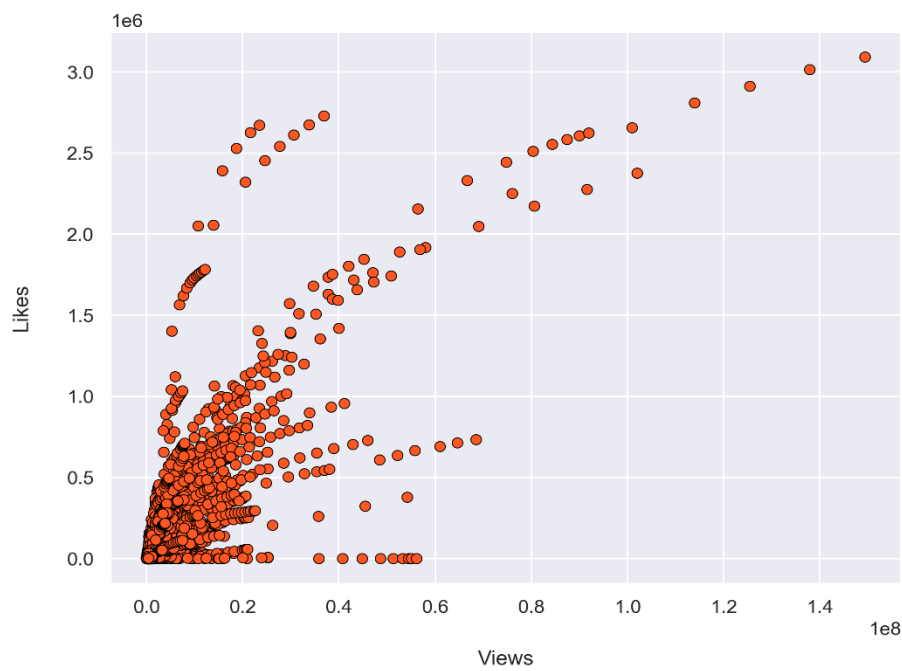
Similarly, we have observed that the percentage of videos with less than 40000 likes is 78.4%. Similarly, we can see that the percentage of videos with less than 100,000 likes is around 89.66%.

Let's observe the correlation:

- To find reasonable results, I removed some categorical columns which are video id, category id, comments disabled, and ratings disabled.
- Thus, we can see the correlation of view count, likes, dislikes, and comment count using a heat map. We can see with the coefficient of 0.79, the highest correlation is between views and likes.



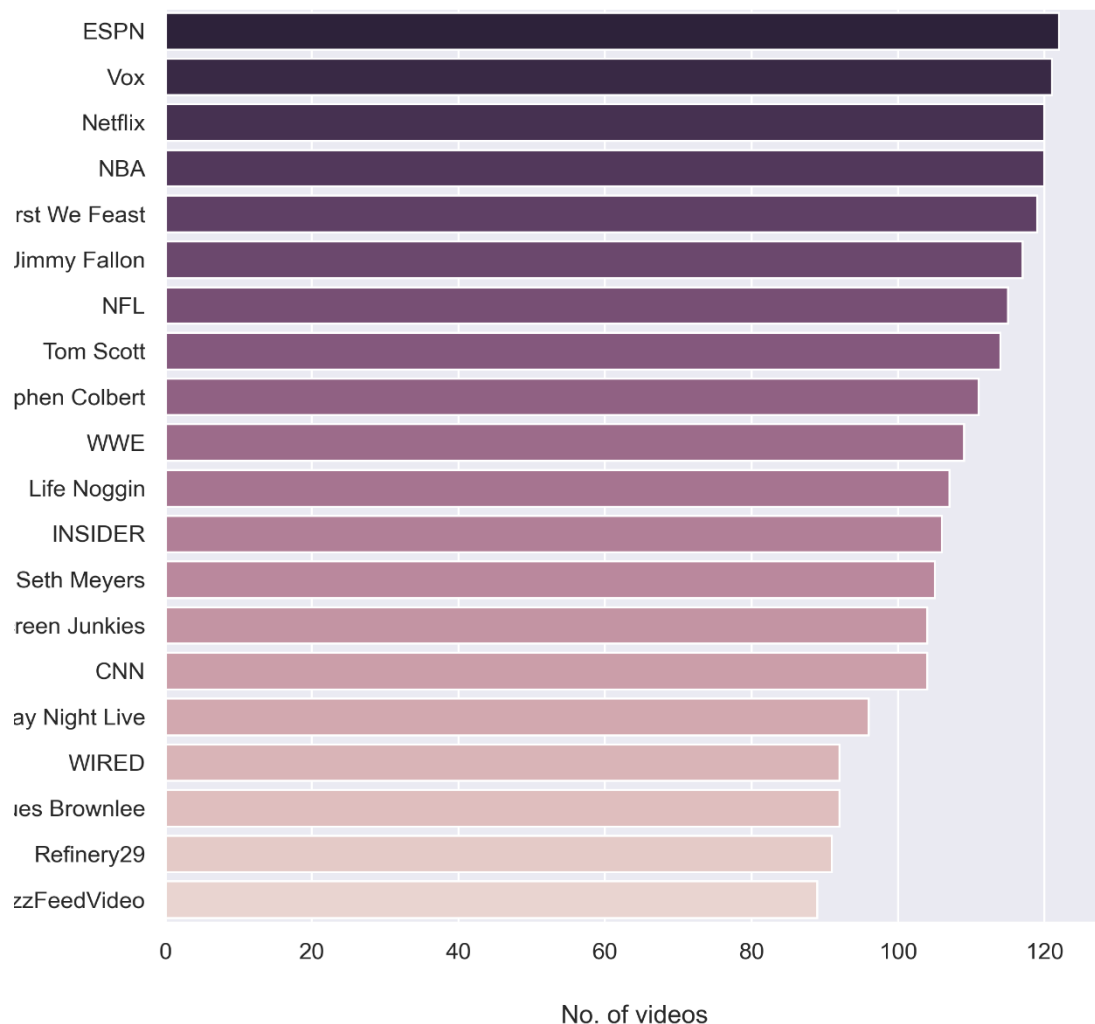
The linearity between views and likes is seen more clearly below:



Views and likes are positively correlated: mostly as one increases, the other increases too.

Trending Channels

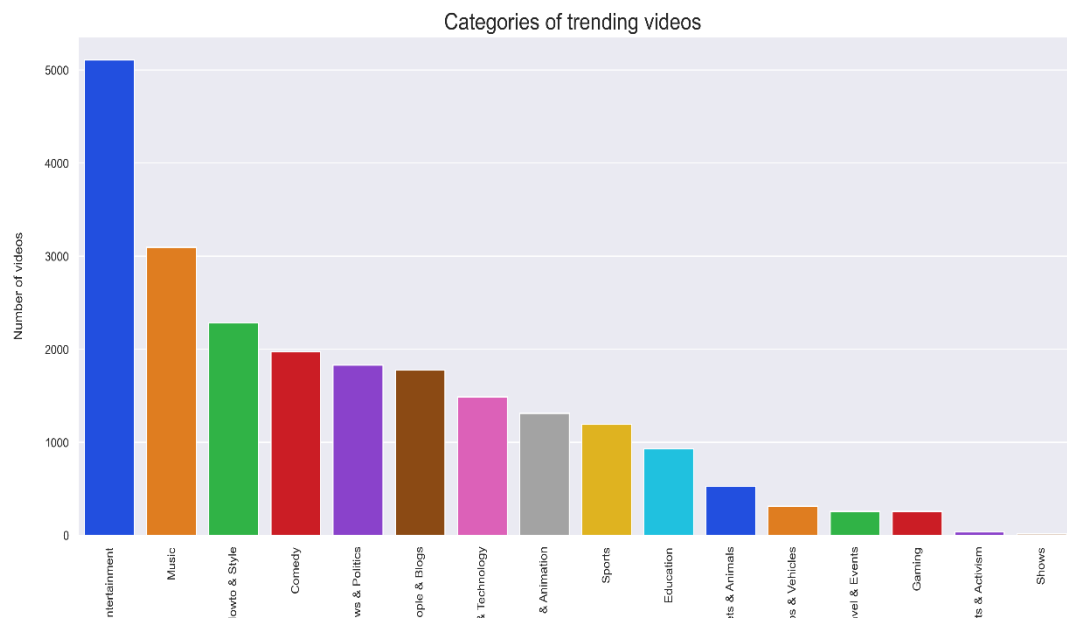
Which channels produced more trending videos? The following bar chart tells us the answer:



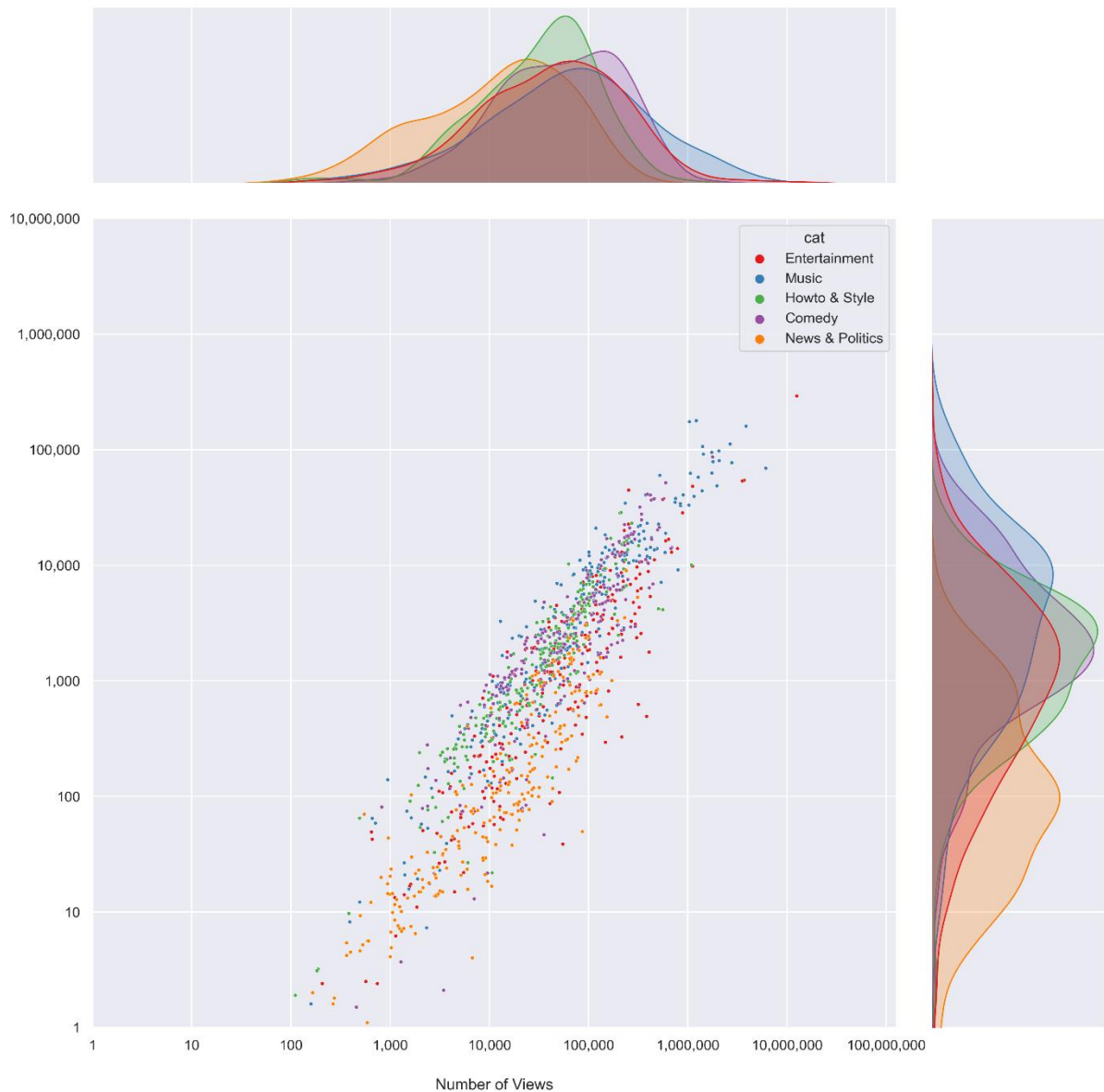
We can see that ESPN is the channel with most trending videos (122 videos) followed by Vox and Netflix.

Trending Videos Categories

The following bar chart shows the number of trending videos for each category.



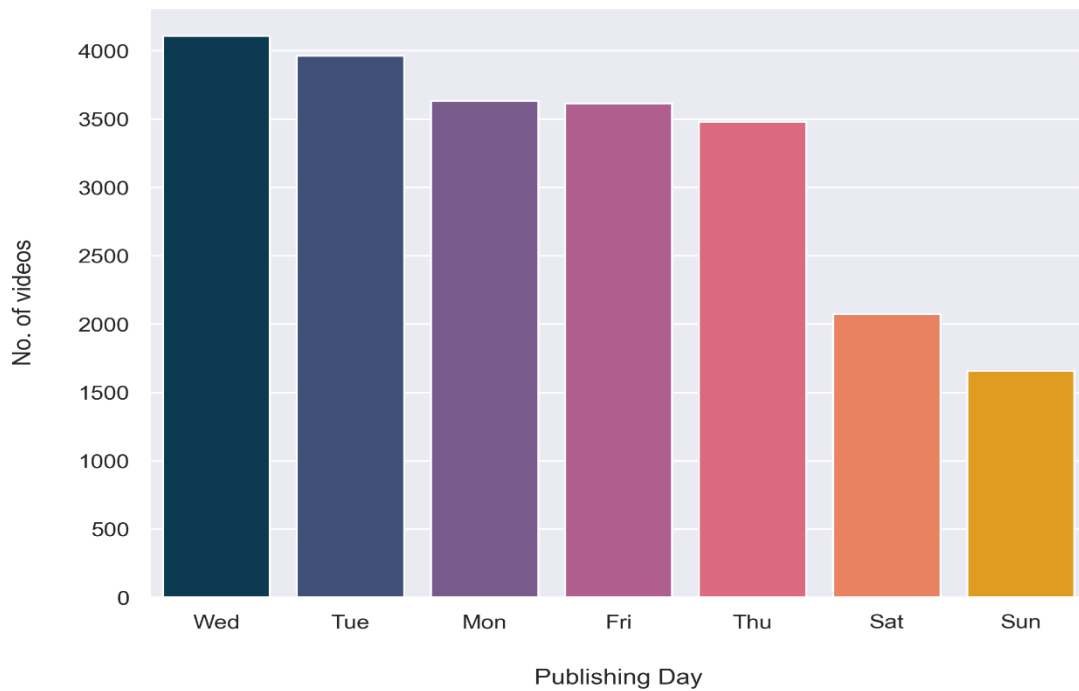
- Entertainment category comes first with 5106 trending videos.
- After Entertainment comes Music category with 3095 trending videos.
- Then at the third place comes Howto & Style category with 2280 trending videos.



As we can see the most viewed and most commented videos are dominated by the musical category, however we are only looking at the top 5 categories by population size.

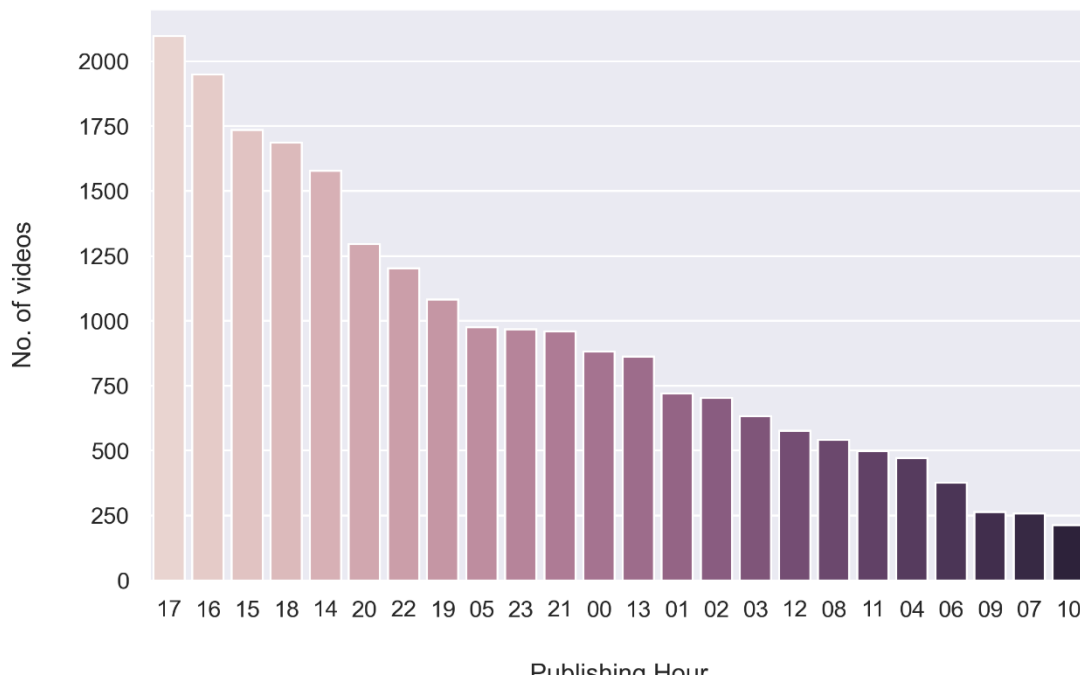
When are the Trending Videos Published?

The following graph shows the number of trending videos published in every day of the week sorted by the number of videos:



- We can see that there were more trending videos published on Wednesday (4109) than the other days of the week.
- All other days of the week except Sunday and Saturday are not far behind of Wednesday.
- For Sunday, there were only 1656 trending videos published on it.

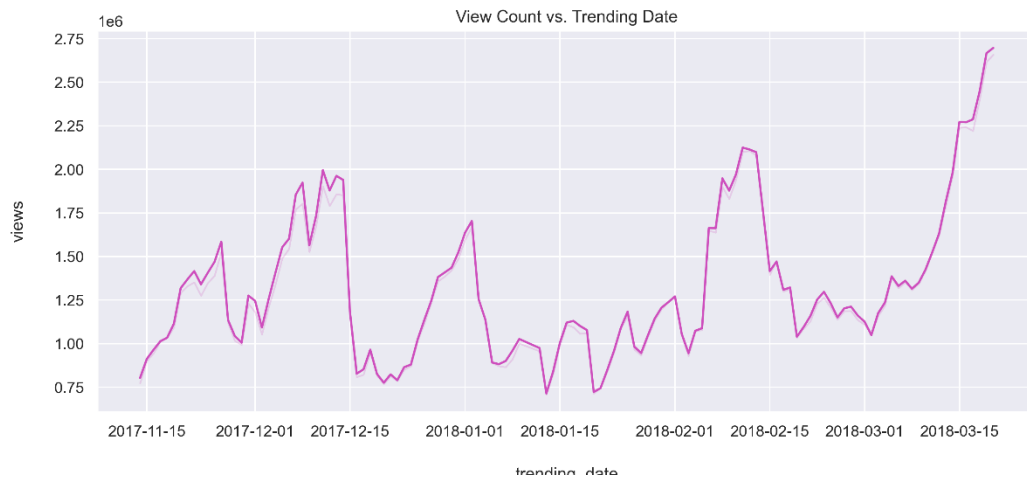
The following graph shows the number of trending videos published in every hour of the day starting with 0 which represents 12 AM and ending with 23 which represents 11 PM:



- We can see that the peak hours for publishing trending videos were between 16 and 18 (between 4 PM and 6 PM), peaking between 4pm and 5pm had largest number of trending videos.
- And the time between 6am-10am has the least number of trending videos with a slight increase at 8am in particular.

The two graphs above don't necessarily mean that videos published on Wednesday have higher chances of becoming trending because Wednesday might be the day that witnessed more uploaded videos in general than other days.

Let's see the relationship between view counts and trending date



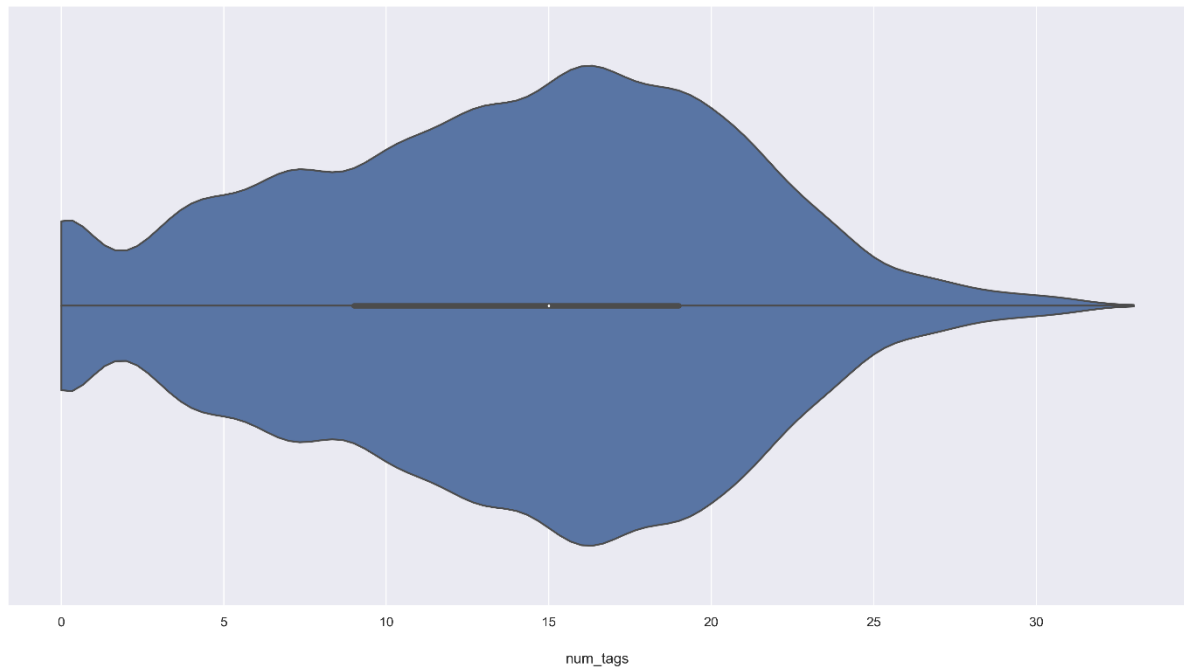
We can see that views are really high after the 15th of March, 2018 compared to all trending dates. And the least views observed around the 15th of January, 2018.

What are the Most Common Words in Video Titles?

Are there some words that occur in trending video titles more than others?

To find out the answer, have analysed the titles of all trending videos and counted the occurrences of each word in those titles.

Here is a word cloud of the most common 100 words in the trending titles. The size of the word reflects how common it is:

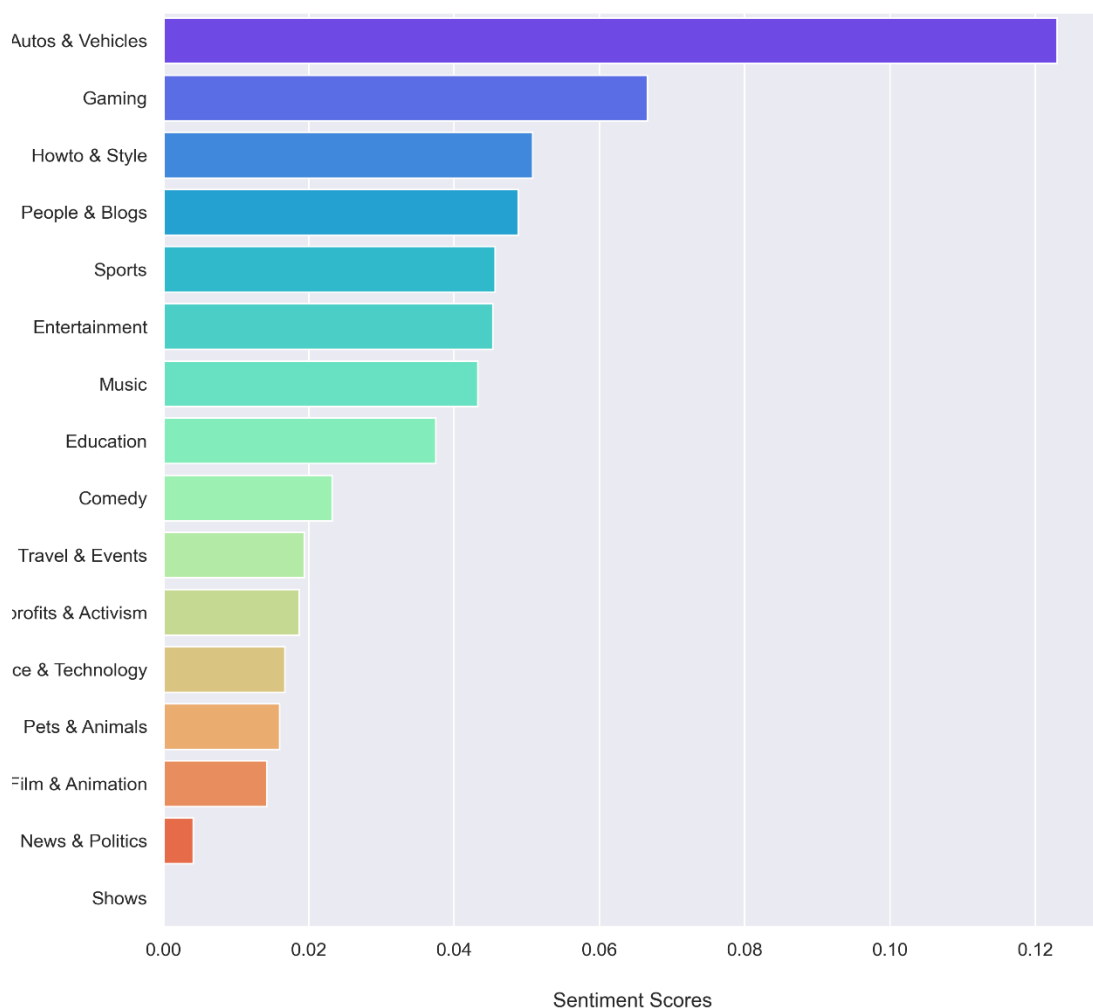


We can see that most of trending videos have between 10 and 20 tags. Here are some interesting facts about trending videos tags:

- The average number of tags a trending video has is **13** tags.
- The minimum number of tags a trending video has is 0.
- The maximum number of tags a trending video has is 33.
- Only 3.5% of trending videos have no tags.

Most Common Words in Trending Videos Tags

The most common tags used in trending videos which are shown in this word cloud:



Tags of videos of certain categories are perhaps unsurprisingly more positive than others, for instance it is intuitive to see why Auto and Vehicles, which often are used to depict a glamorous view would have a more positive description than News and Politics which often describe a story in an unfavourable way depending on the political affiliation of the channel and their respective values align with the story they are describing.