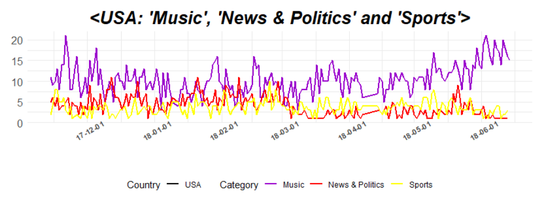
YouTube top trending videos

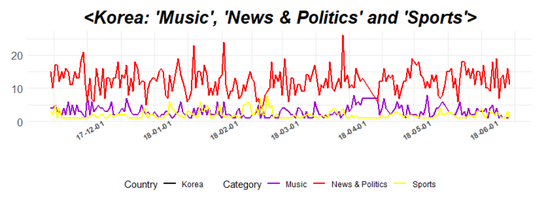
2015-16506
Architectural Engineering
LEE HyoBae

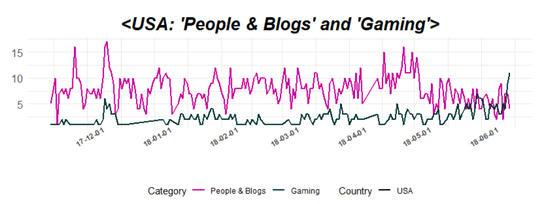
YouTube data

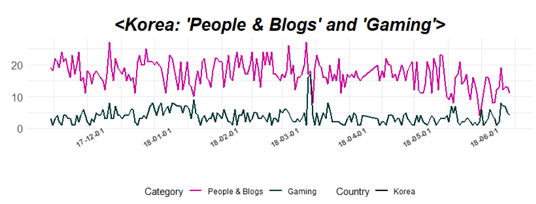
YouTube is globally used and data stacked inside the platform reflects trend and issue of world. I utilize data of top trending videos in YouTube. It consists of 10 countries(Canada, Germany, France, UK, India, Japan, South Korea, Russia, Mexico, USA) and numerous features including numerical indicators('views / likes / dislikes / comments').

Category's Trend Line



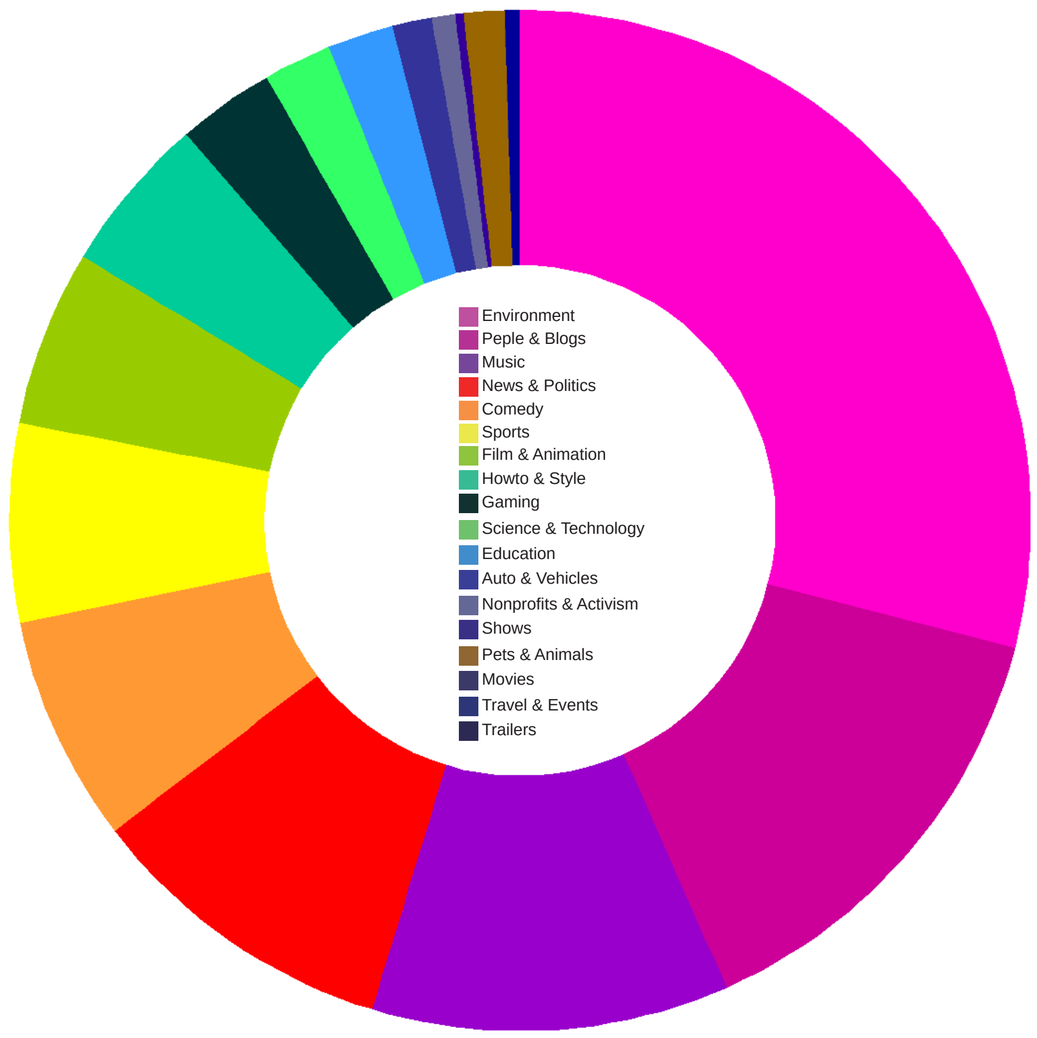




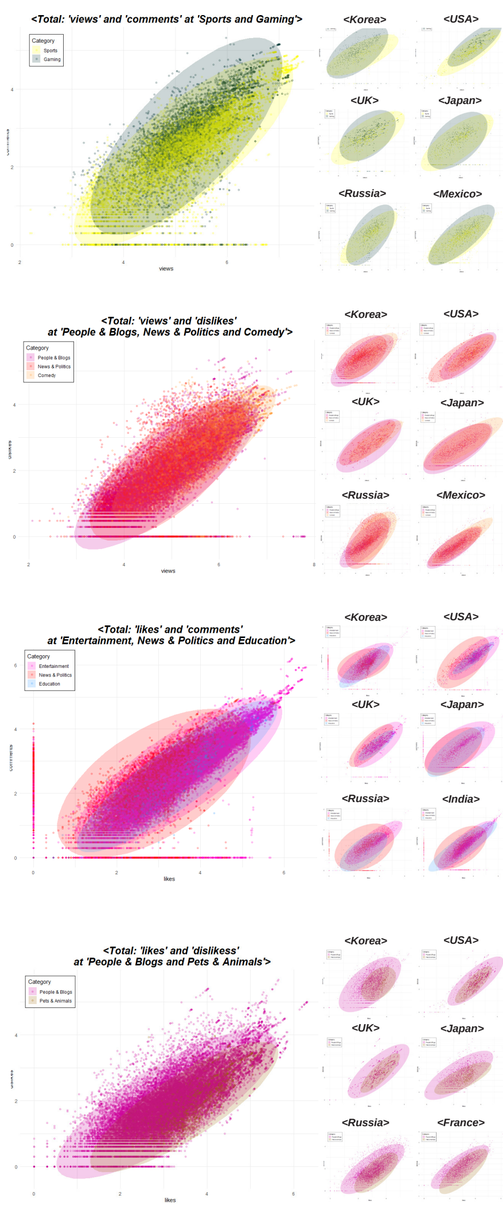


Category Proportion

There are 18 categories and proportion of each category is different for each country. For instance, 'News & Politics' takes quite huge portion in South Korea, unlike other countries. I draw the trend line of each category per countries and figure out the difference among countries. Also I find out that for each category, the relation between views/likes/dislikes and comments differs. Moreover, the colony's shape of category varies among countries.



Category's Scatter Plot



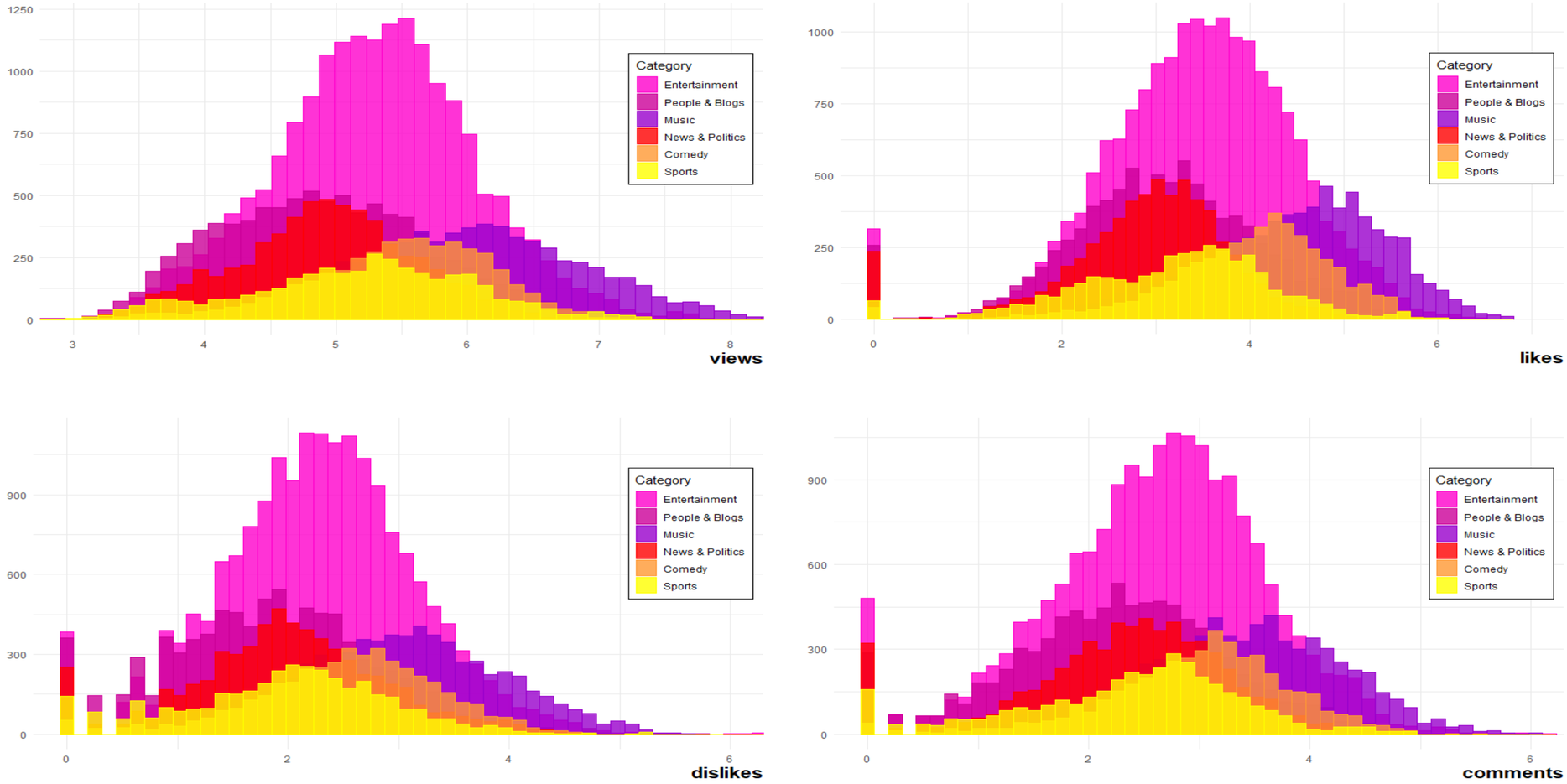
Category's Histogram

For those 6 of largest number of viedeos, I draw histograms of categories for 'views / likes / dislikes / comments' and examine distribution.

As you can see, 'Music' tends to get high numbers at all numerical indicators while, 'Peple & Blogs' and 'News & Politics' do not. One of the reasons is that the portion at which a video disable rating and commenting is higher in 'People & Blogs' and 'News & Politics'.

Another interesting point is that distributions of 'vews / likes / dislikes / comments' are all similar. Before I look into data, I expect that at 'dislikes' some categories(e.g. 'News & Politics') would be more to the right. However, The histogram shows that the videos which are viewd by many users also get large numbers of 'likes', 'dislikes' and 'comments'.

To look at histograms for each country. A remarkable point is that although the proportion of each categories are different among countries, dispostion of each histogram is similar to each other. For example, although 'Music' takes small part in South Korea, the histogram of it places more right than other categories for all numerical indicators.

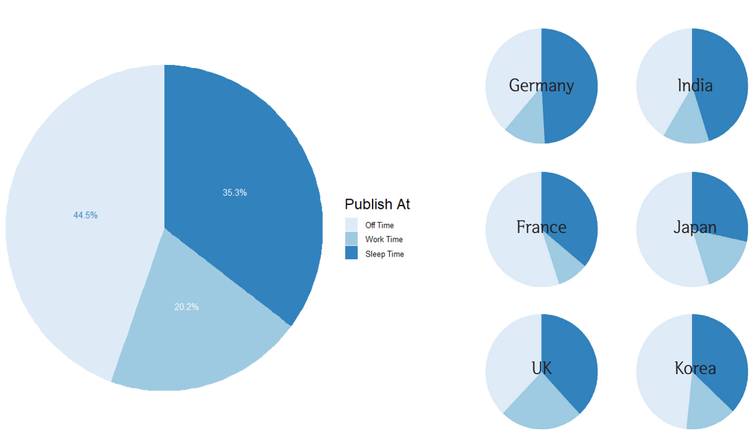


Word Cloud of Tags

Using tags of each video, I make wordclouds for USA and South Korea. An interesting point is that the tags which appear most is totally different between two countries. While USA videos' most popular tags are 'funny' and 'comedy', South Korea's are president's name and 'mukbang'.

Time video uploaded

Most of the videos are uploaded after the quitting time becuase it is the time most people uses YouTube. I divide a day into three part (Sleep Time, Work Time, Off Time) and examine the difference of uploading tie among countires. Interestingly, while in some countries(France, Japan), most videos are uploaded during 'Off Time', videos are uploaded at more various time in others.



Disable rating/commenting

Some videos do not allow users to like, dislike or comment. The reason is because of the video's controversy or preventing malicious feedbacks. The proportion of disableing status is different for each category.