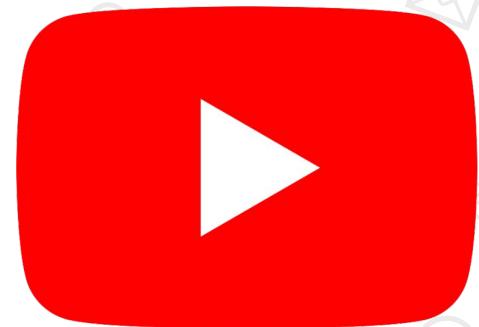


YouTube Trending Videos

ECON 4590 - Chuheng Yu



YouTube? Trending?

- Video sharing platform
- Created on 2/14/2005
- 720,000 hours of video uploaded in a day
- Trending launched in May
- 50 videos on the chart
- Updated every ~15 mins
- Not personalized
- Remain the same for each country

Trending videos

- Chelsea vs. Real Madrid: Extended Highlights | UCL Quarter-Finals - Leg 1 | CBS Sports Golazo**
CBS Sports Golazo 1.1M views • 1 day ago
Extended Highlights, Presented By Heineken Christian Pulisic rejoins Chelsea after a hatrick on international duty to host Champions League legend Karim Benzema and Real Madrid. Stream every...
- Nicki Minaj Carpool Karaoke**
The Late Show with James Corden 1.9M views • 17 hours ago
James and Nicki Minaj cruise through Los Angeles to her biggest hits from "Starships" to "Super Bass" to "Do We Have a Problem?" And after flexing her incredible British accent, we get the...
- AMP LOVE IS BLIND**
AMP 569K views • 1 day ago
Join this channel to get access to perks: <https://www.youtube.com/channel/UCJbYdyufHR-cxOuY96KloqA/join>
AMP AMP AMP AMP 💕 Join the AMP SQUAD on Discord: <https://discord.gg/zMeuncfUcB> ...
- Josh Brolin Licks the Palate of Absurdity While Eating Spicy Wings | Hot Ones**
First We Feast 517K views • 7 hours ago
Josh Brolin is an Academy Award-nominated actor with storied credits that include No Country For Old Men, Sicario, Deadpool, The Avengers, and many more. His latest project is the Western-set...
- Marvel Studios' Doctor Strange in the Multiverse of Madness | Dream**
Marvel Entertainment 5.7M views • 1 day ago
The Madness begins 🎟 Tickets are now on sale for Marvel Studios' #DoctorStrange in the Multiverse of Madness. Experience it only in theaters May 6. Get Tickets Now: fandango.co/3NQZxgH...



Determining Factors

What Are the Factors:

- Appealing to wide range of viewers
- Not misleading, clickbaity or sensational
- Diversity
 - Videos
 - Creators
- Surprising/Novel

How To Measure:

- View count
- “Temperature”
- Where views are coming from?
- Age of the video
- Compared with other video in the same channel



“We combine these signals to produce a list of videos that showcases what’s happening on YouTube.”

-- By YouTube Help Center

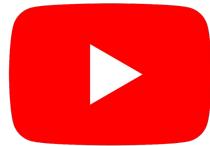




Literature Review

Understanding Digital Ethnography: Socio-computational Analysis of Trending YouTube Videos -- Published in 2018

- Data:
 - Top 200 daily trending videos from 9/2017 to 10/2017 in USA and GB
 - Attributes: *URL, Video ID, Video Title, Channel Title, Category, Views #, Likes #, Dislikes #, Comments #*
- Collection phase:
 - Obtained from Kaggle
 - Enhanced the dataset by using YouTube API: *Description, Video Created Date, Subscribers #*
 - Collected social media associations using Web Content Extractor



Key Findings

- "More social media presence you have, the more the views." ✗
- Videos in the Music category were shared more to different social media sites

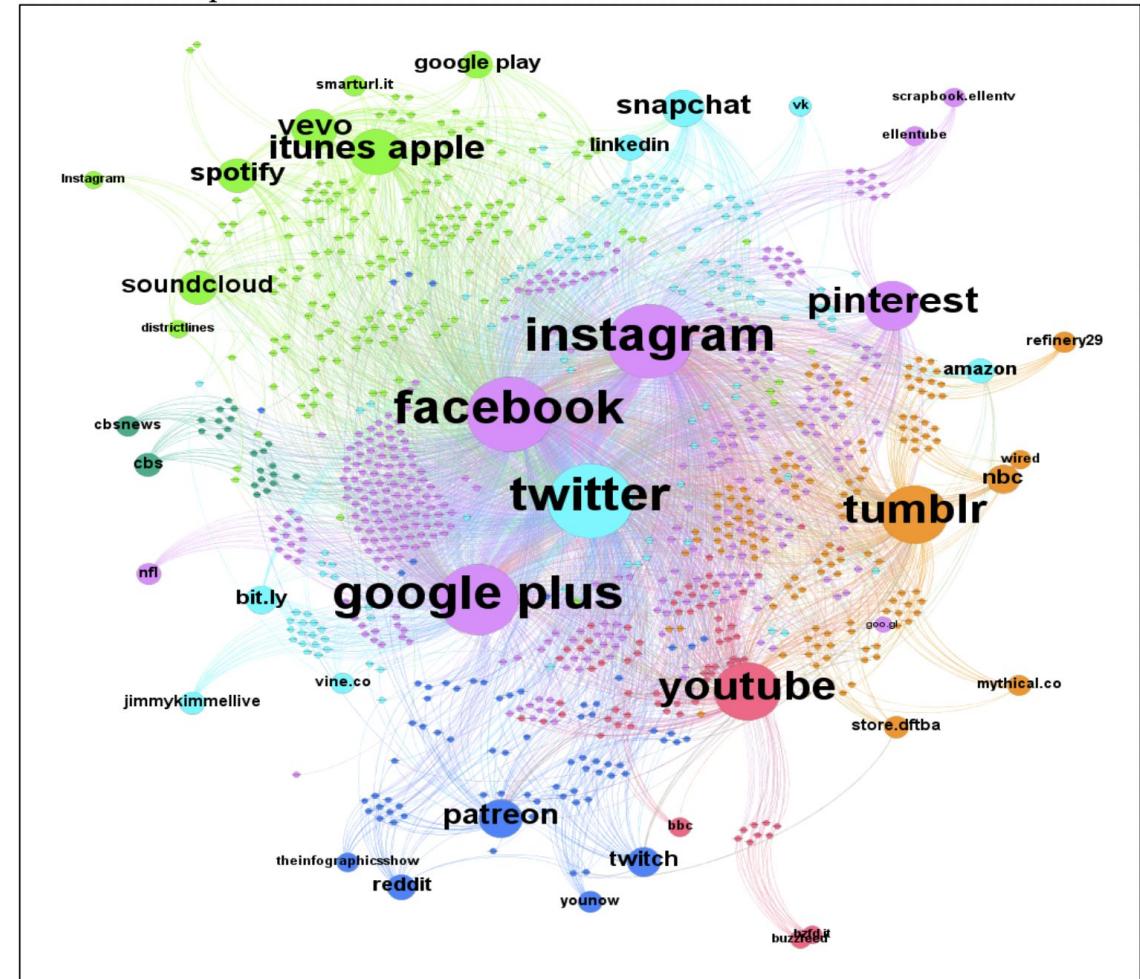


Fig. 1. Social media map of the trending videos in USA.

Key Findings

- More comments in a video, the more neutral the sentiments.
- More likes -> more neutral the description sentiment
- Strong correlations between:
 - Views# & Likes#
 - Comments# & Dislikes#
- Weaker correlation between:
 - All other categories

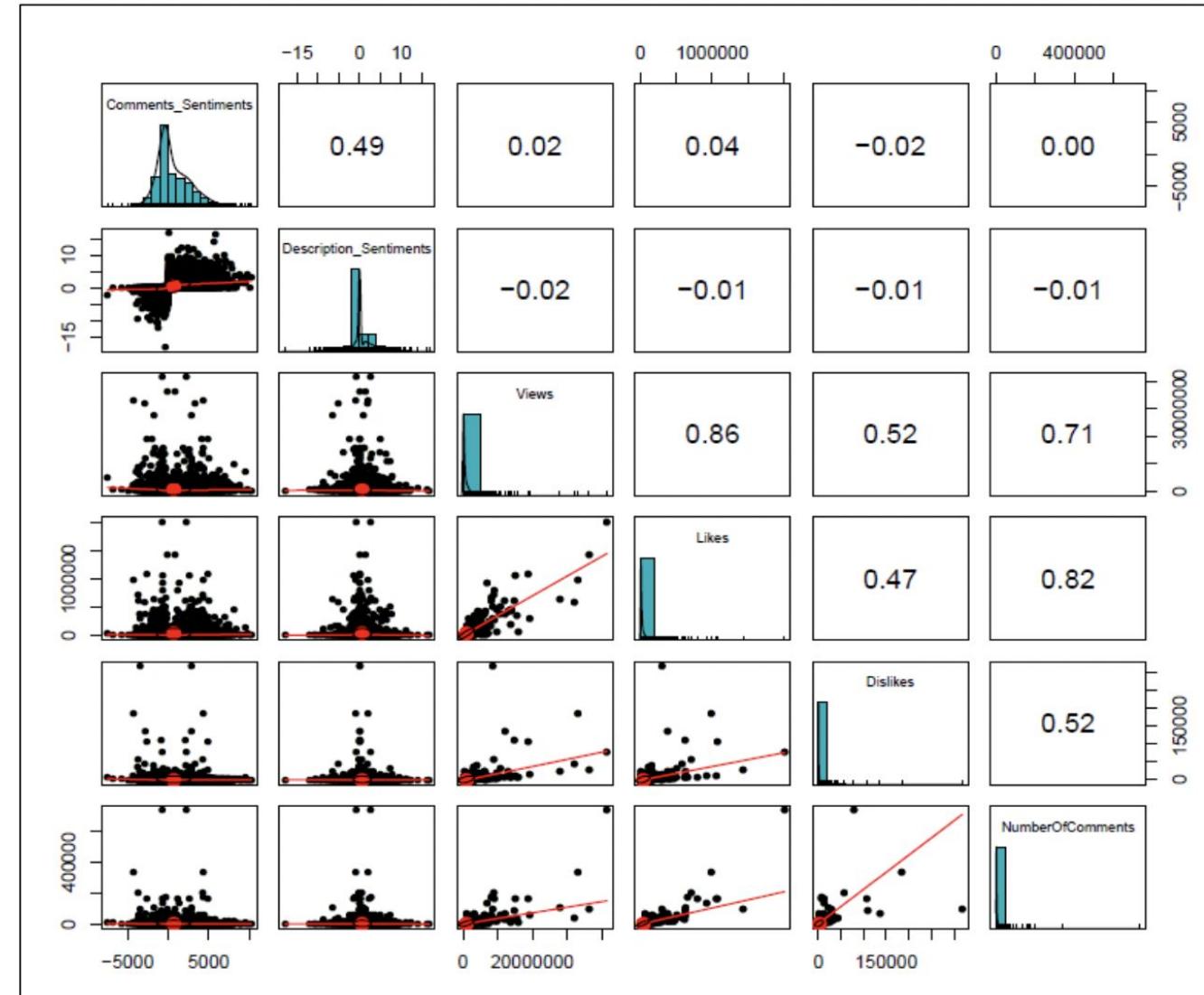


Fig. 2. Correlation analysis of the trending videos in USA.



Conclusions

- Likes # and Views # have a strong positive correlation as comparing to correlation between Dislikes # and Views #
- Fewer comments a video has the more polarized the sentiments are
- If a viewer watches the whole video, he is more likely to like the video than dislike it
- Commenting behavior is somewhat unrelated to the content



Literature Review

Interactions between emotional and cognitive engagement with science on YouTube --

Published in 2021

- Focus on:
 - What is the proportion of engagement with science within top trending YouTube videos?
 - What characterizes behavioral, cognitive, and emotional engagement with scientific content?



Methods

- Data Collection
 - Used YouTube API
 - 3-month period from March 2019
 - 200 trending videos a day
- Behavioral engagement with science content
 - Focused on Science and Education category
 - # of Views, Likes, Dislikes, Comments
 - Random sample of 1000 comments per video
 - One-way analysis of variance(ANOVA)
 - Used Syuzhet R package
 - Anger, anticipation, disgust, fear, joy, sadness, surprise, trust
 - Positive, negative



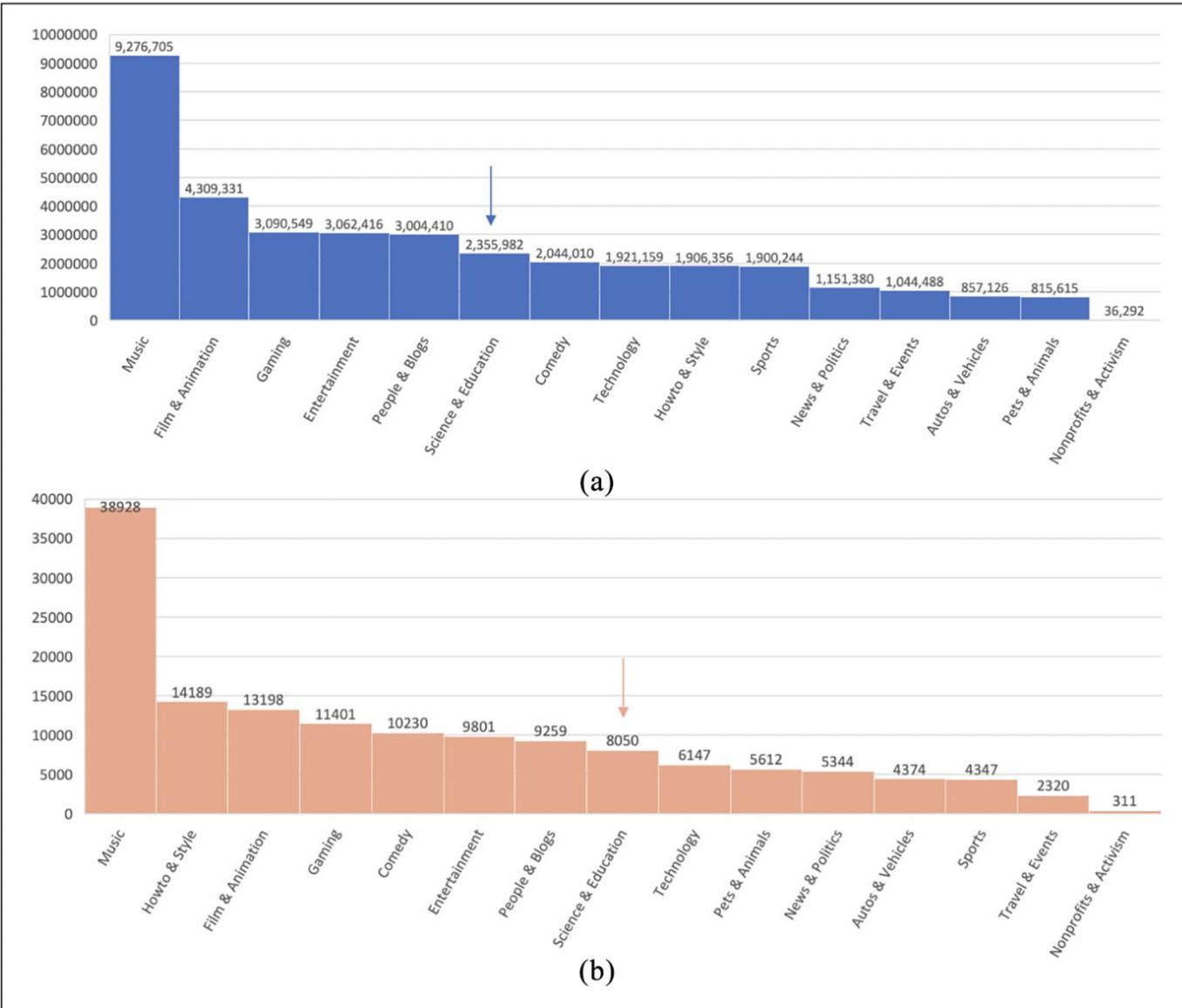


Figure I. (a) mean number of views per category and (b) Mean number of post-video comments per category.

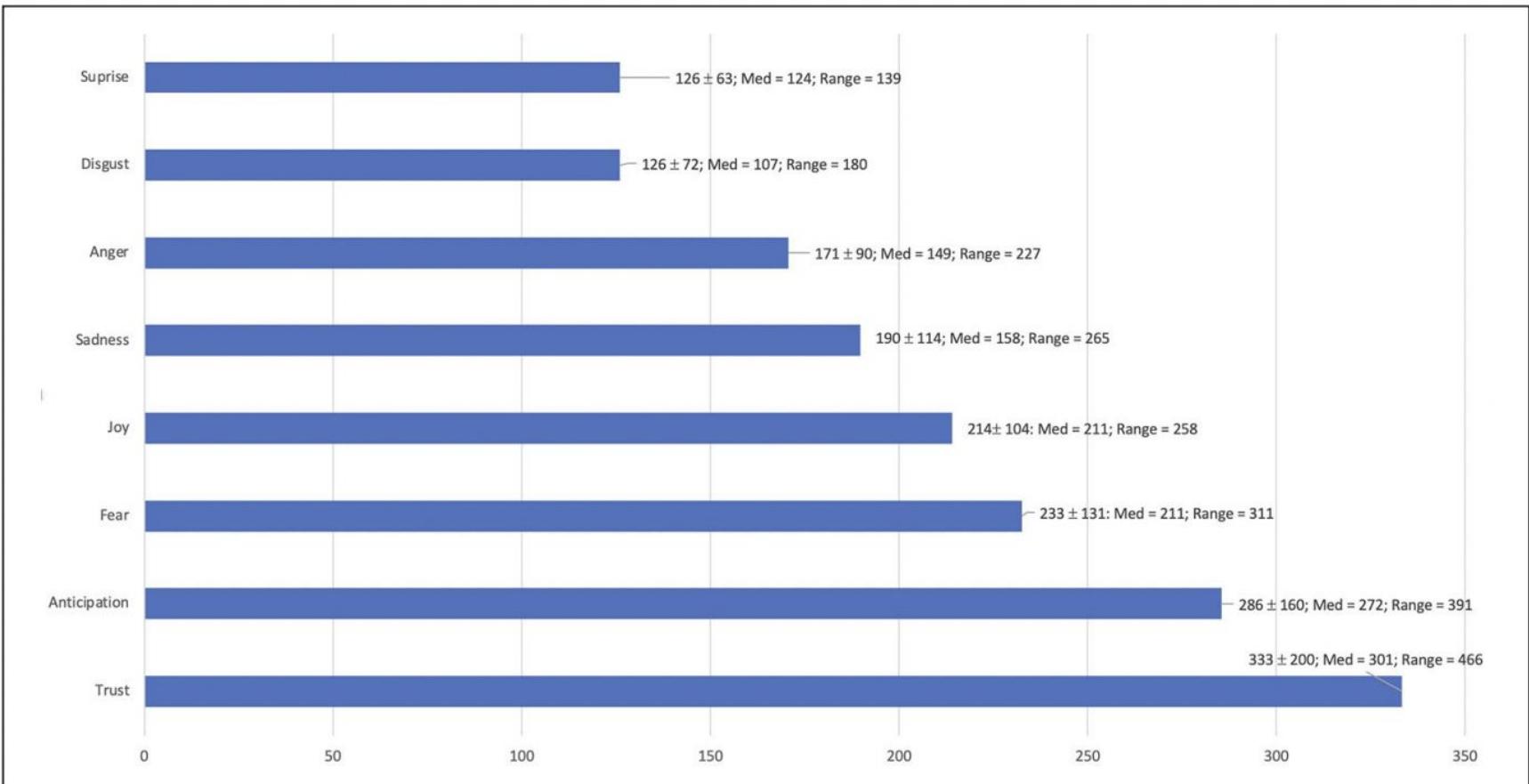


Figure 2. Sentiment analysis: Amount of emotional expressions per video (Mean \pm SD; median; 10–90 percentile range).



Key Findings

- Top 3 categories:
 - Entertainment (30%)
 - Music (14.5%)
 - Sports (11.4%)
- Science & Education
 - Only 3.4%
 - High average number of views
 - High number of post-video comments on average
- Views#
 - 20 times more than Likes#
 - 300 times more than Comments#
 - Positive & negative emotions are both associated with comment engagement



Current Process

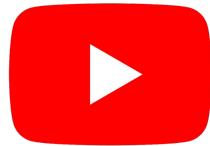
- model <- comment_count ~ category_id + views + likes + dislikes + titlecount + tag_count
 - Comment participation from different categories in trending videos
 - Which factor has a larger effect towards the result?
- Combined all datasets into one dataframe
 - Macro perspective
- Created new columns
 - title_count
 - country
 - tag_count



Stranger Things 4 | Official Trailer | Netflix

Stranger Things 4.6M views • 15 hours ago

It's time. See you on the other side. Stranger Things 4 Vol. 1 premieres May 27th, only on Netflix.



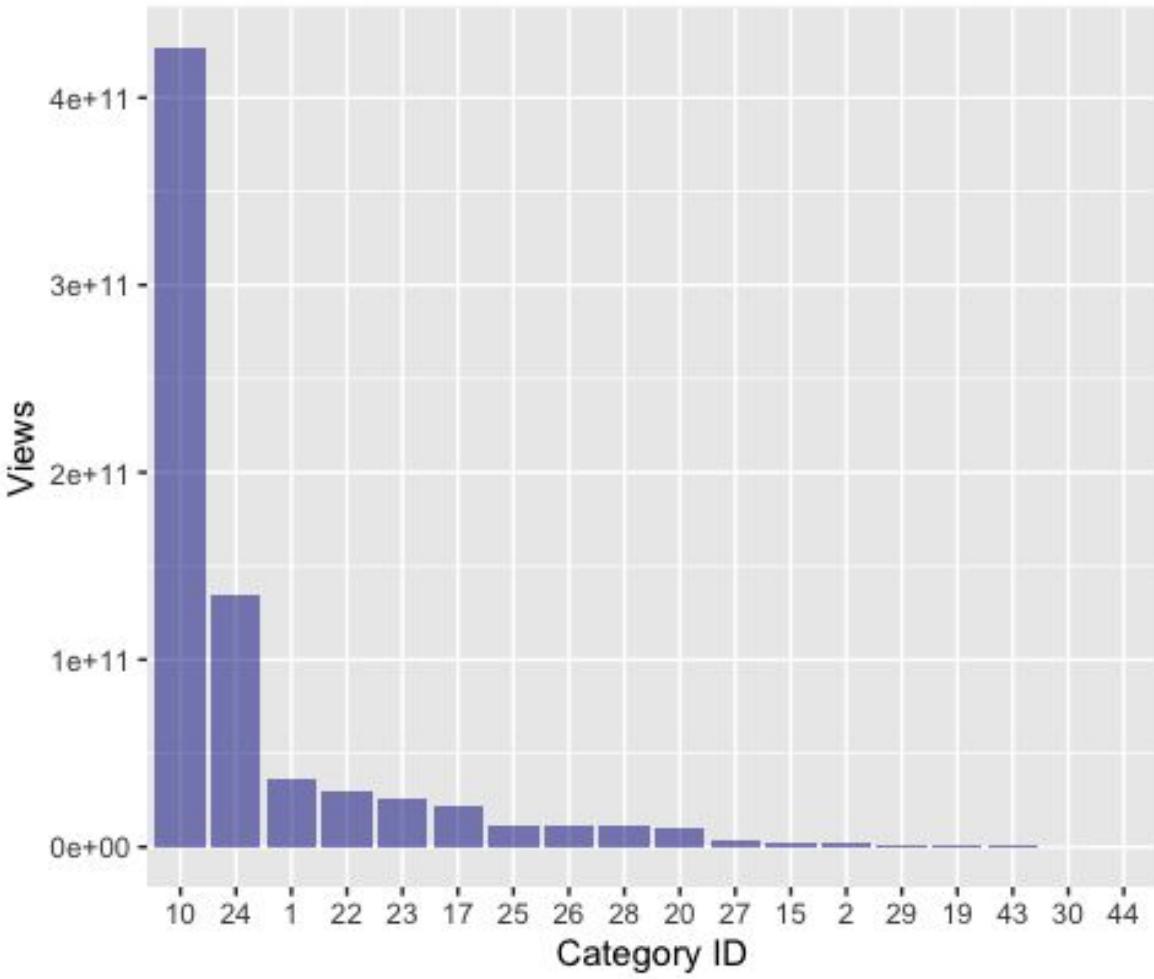
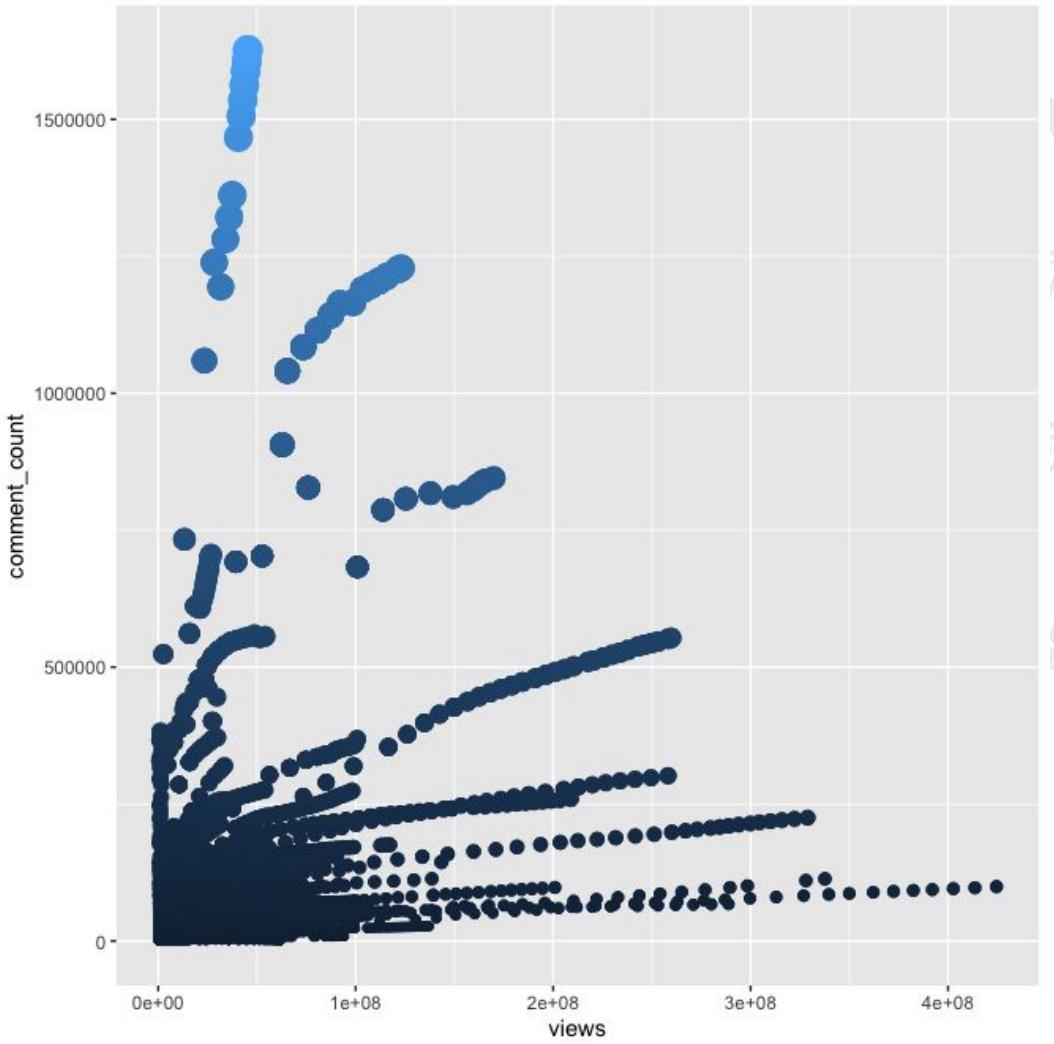
Dataset

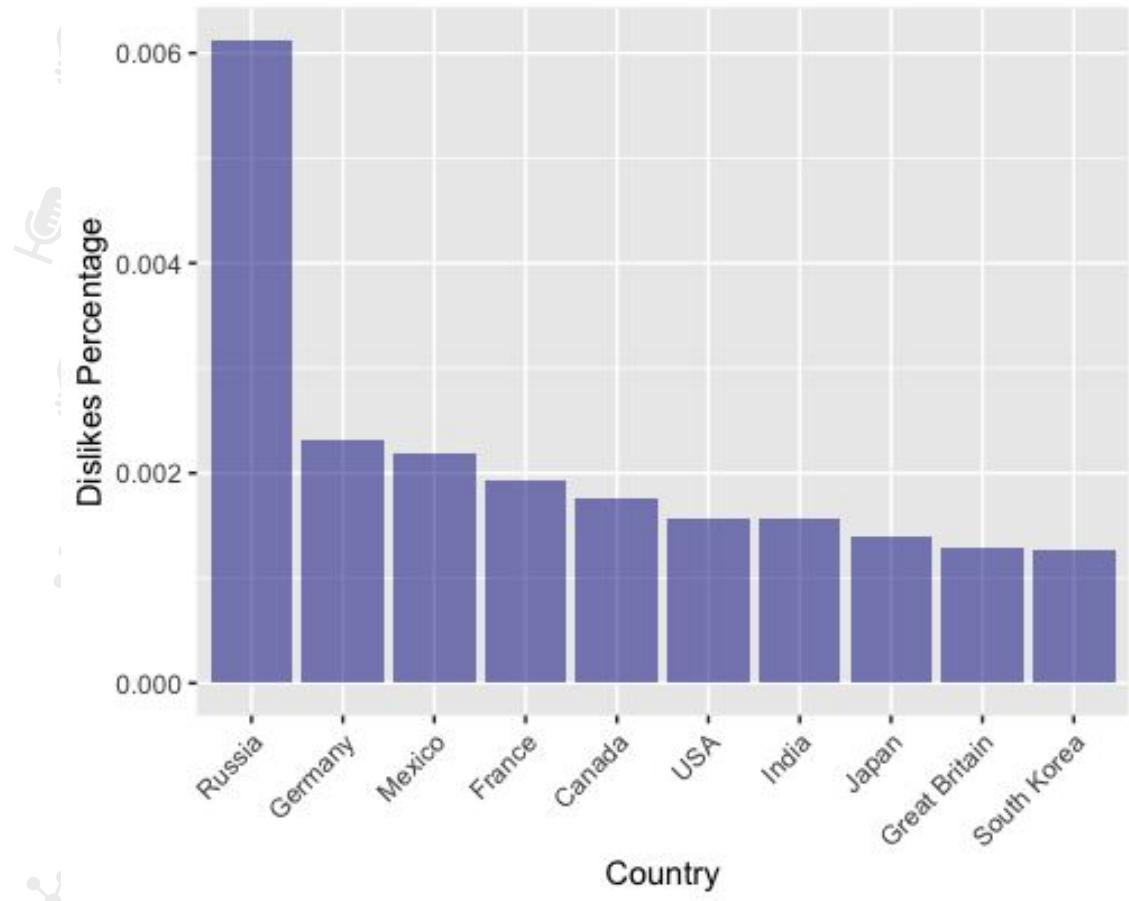
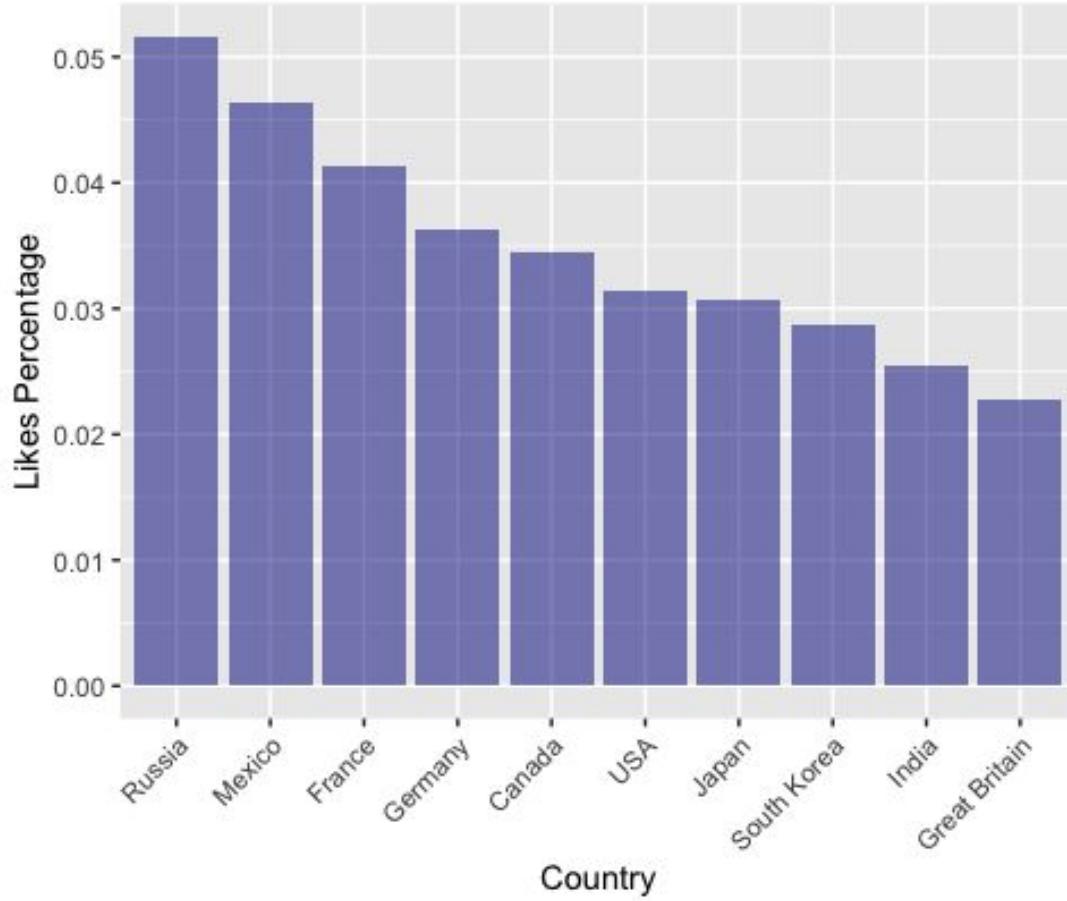
Dataset:

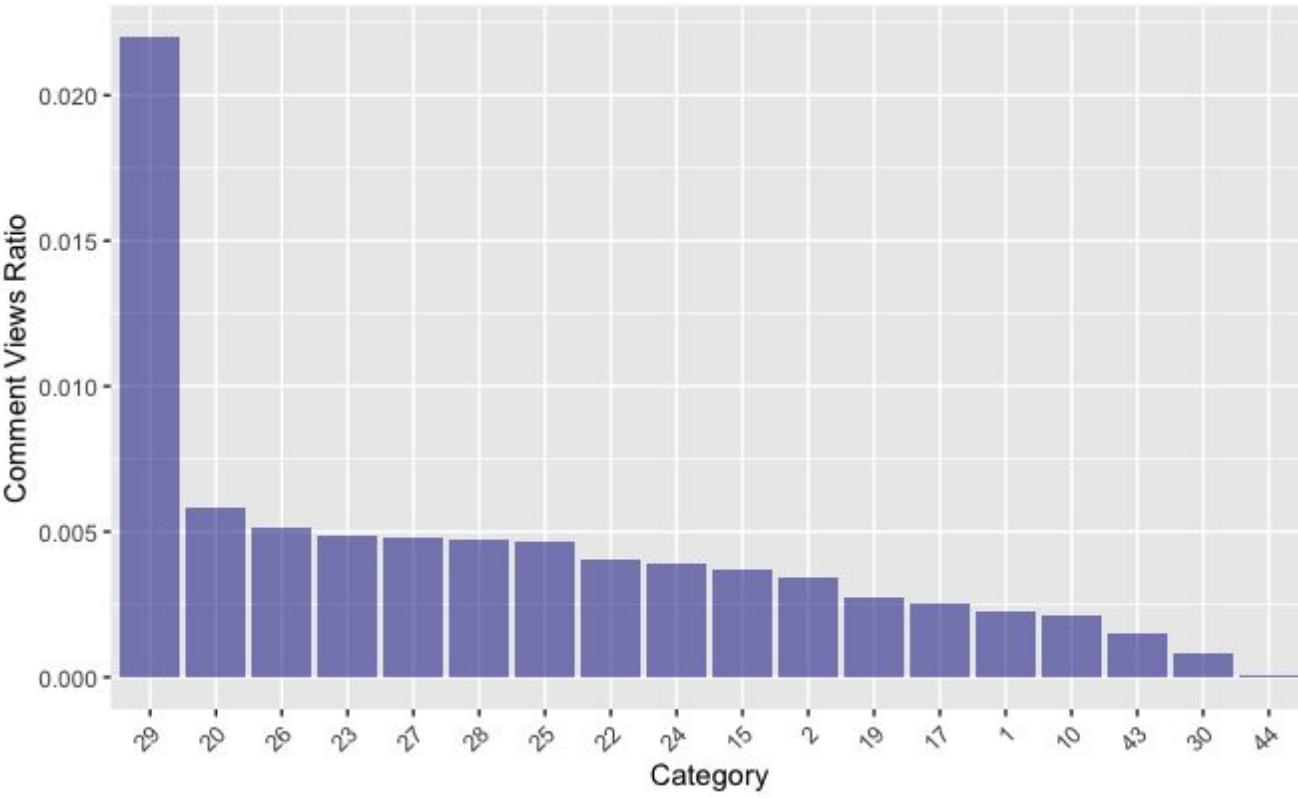
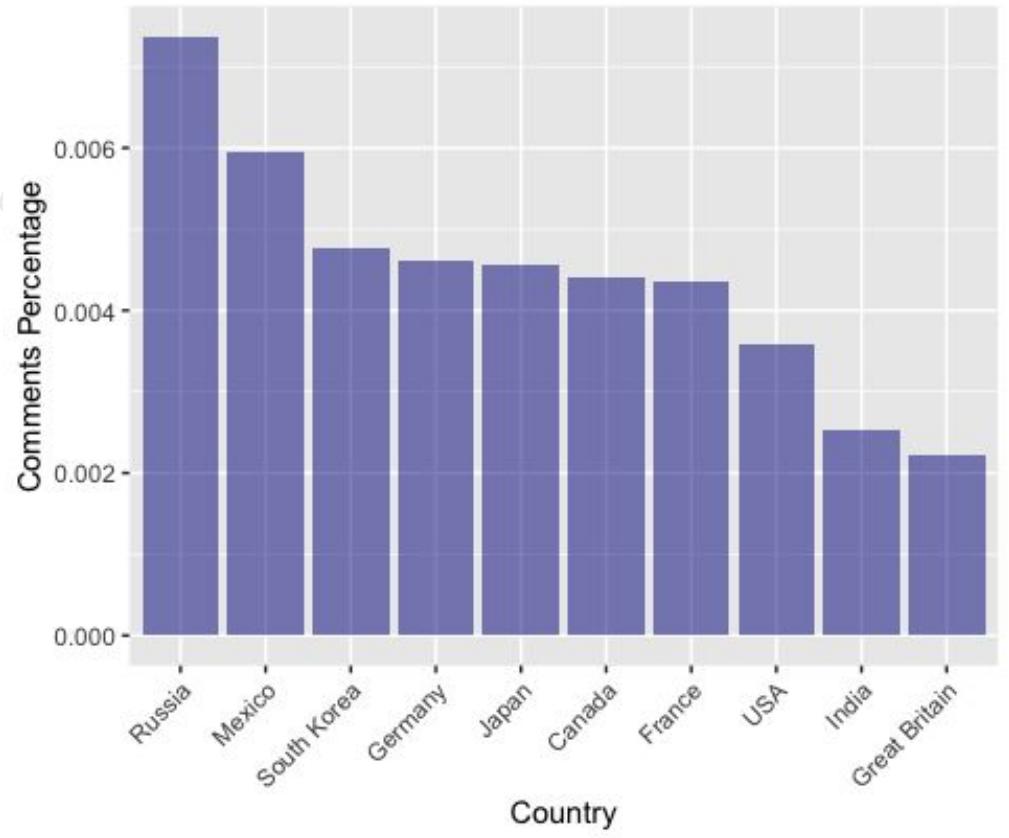
- Russia, Mexico, South Korea, Japan, India, USA, Great Britain, Germany, Canada, and France
- 200 videos per day
- From 12/1/2017 to 5/31/2018
- 414,858 observations

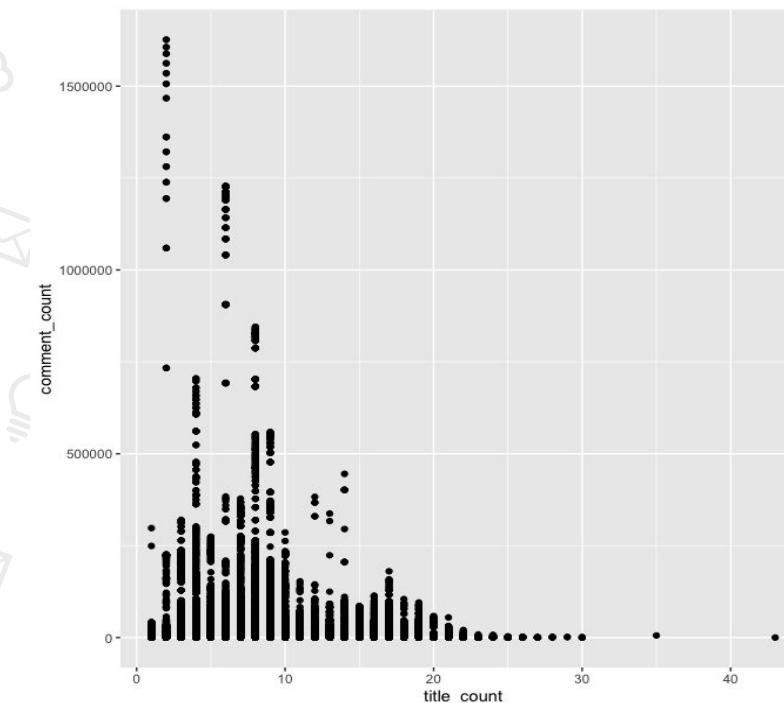
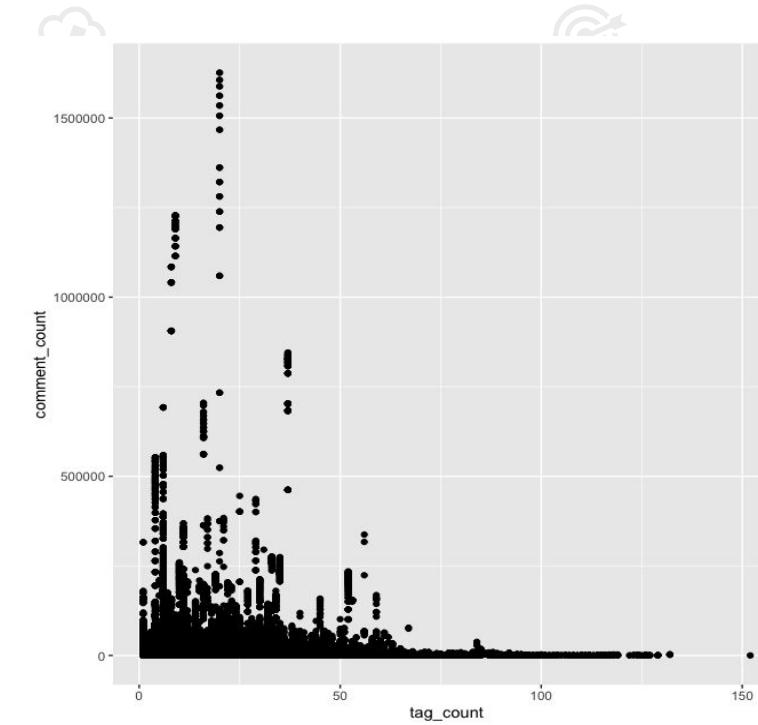
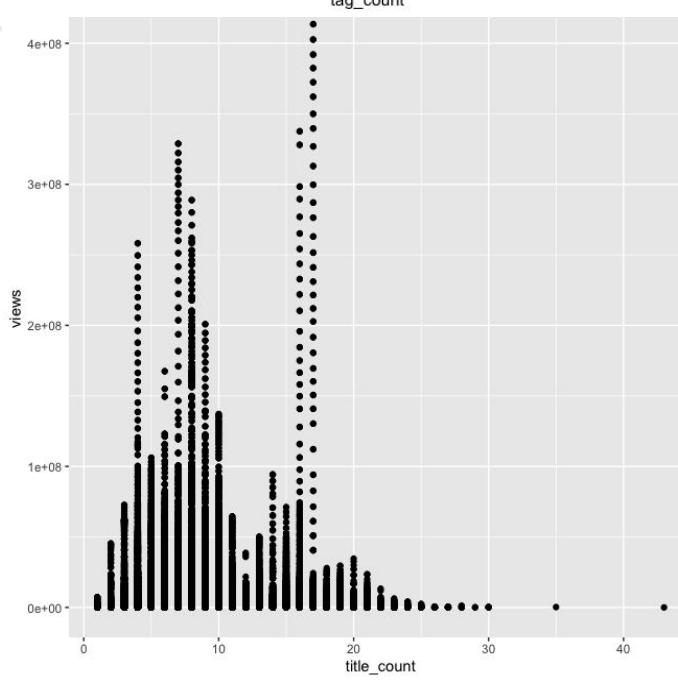
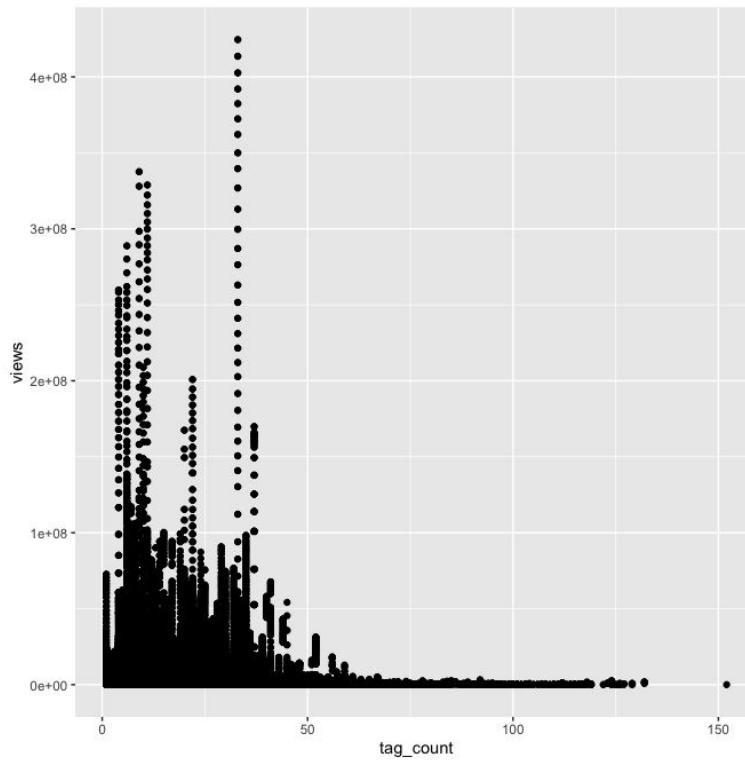
Attributes:

- video_id, trending_date, title, channel_title, category_id, publish_time, tags, views, likes, dislikes, comment_count, thumbnail_link, comments_disabled, ratings_disabled, video_error_or_removed, description,









OLS:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	2.1649e+05	3.2396e+04	6.6828	2.746e-11 ***
views	-1.2041e-03	2.2750e-04	-5.2928	1.284e-07 ***
likes	1.2917e-01	1.3482e-02	9.5811	< 2.2e-16 ***
dislikes	5.1480e-01	3.6546e-02	14.0862	< 2.2e-16 ***
title_count	-1.2185e+02	4.7543e+01	-2.5629	0.0104253 *
tag_count	-1.7930e+01	2.6399e+01	-0.6792	0.4970540
category_id_2	-1.7902e+05	2.5519e+04	-7.0151	2.781e-12 ***
category_id_10	-9.2687e+05	1.8975e+05	-4.8846	1.086e-06 ***
category_id_15	-1.7344e+05	2.5301e+04	-6.8550	8.490e-12 ***
category_id_17	7.2064e+03	1.5744e+04	0.4577	0.6471836
category_id_19	-1.9688e+05	2.9338e+04	-6.7108	2.273e-11 ***
category_id_20	-9.2276e+04	2.1113e+04	-4.3705	1.278e-05 ***
category_id_22	1.9927e+05	4.4168e+04	4.5115	6.664e-06 ***
category_id_23	-1.4230e+05	4.0099e+04	-3.5487	0.0003925 ***
category_id_24	8.8856e+05	1.4571e+05	6.0983	1.197e-09 ***
category_id_25	1.9551e+05	3.2316e+04	6.0500	1.613e-09 ***
category_id_26	-1.6039e+04	1.2979e+04	-1.2357	0.2166640
category_id_27	-1.6099e+05	2.1021e+04	-7.6589	2.455e-14 ***
category_id_28	-6.3788e+04	2.6441e+04	-2.4125	0.0159001 *
category_id_29	-1.7041e+05	3.2866e+04	-5.1850	2.291e-07 ***
category_id_30	-2.1331e+05	3.1631e+04	-6.7435	1.821e-11 ***
category_id_43	-2.0901e+05	3.0965e+04	-6.7500	1.742e-11 ***
category_id_44	-2.1588e+05	3.2322e+04	-6.6791	2.815e-11 ***

FE:

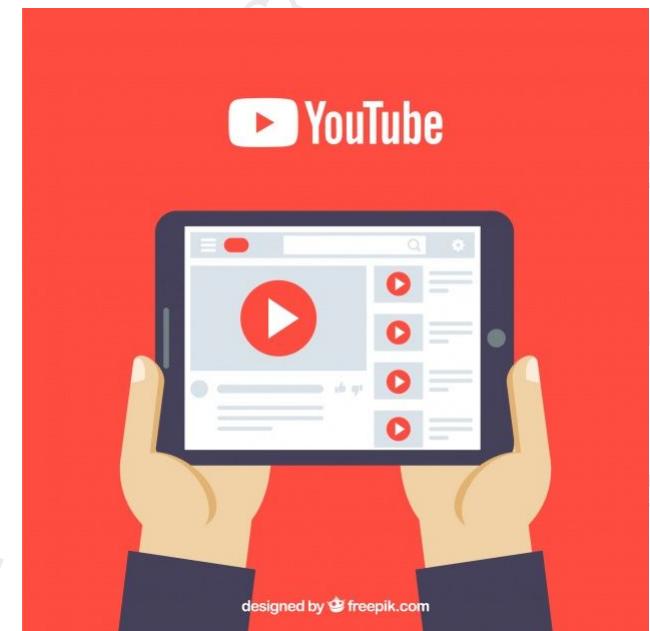
	Estimate	Std. Error	t value	Pr(> t)
views	-1.2041e-03	7.3453e-05	-16.3928	<2e-16 ***
likes	1.2917e-01	2.8392e-03	45.4965	<2e-16 ***
dislikes	5.1480e-01	1.6131e-02	31.9144	<2e-16 ***
title_count	-1.2185e+02	9.2744e+01	-1.3138	0.1890
tag_count	-1.7930e+01	4.0384e+01	-0.4440	0.6571





Results

- Dislike has the highest positive impact on comment participation
- “More views in a video, the more comments” **X**
- Views# almost have no impact on comment#
- When viewer click a like or dislike
 - They are more likely to engage in the comment section
- It less likely to have a high comment participation when:
 - As title gets longer
 - As number of tags increase



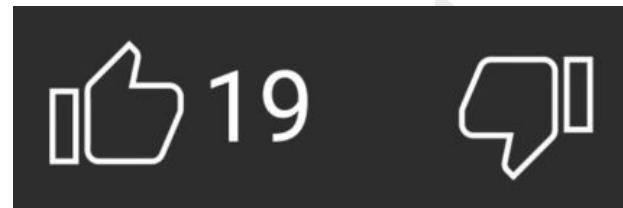


Future Works & Concerns

- Future Works:
 - Focus on a category or creator
 - Geographic (Russia?)
 - Comments Sentiment (Syuzhet Package)

- Concerns:

- Selection Bias?
- Old dataset
- Unpredictable Hypes
- Removal of public dislikes counter
- Changing of names -> “Explore”
- “Shorts” video form



THANK YOU!

Questions?



Citations

- Hussain, M. N., Tokdemir, S., Al-khateeb, S., Bandeli, K. K., & Agarwal, N. (n.d.). *Understanding digital ethnography: Socio ... - sbp-brims.org*. Retrieved April 5, 2022, from http://sbp-brims.org/2018/proceedings/papers/latebreaking_papers/LB_14.pdf
- Dubovi, Ilana, and Iris Tabak. "Interactions between Emotional and Cognitive Engagement with Science on YouTube." *Public Understanding of Science* 30, no. 6 (August 2021): 759–76. <https://doi.org/10.1177/0963662521990848>
- Khanam, S., Tanweer, S., & Khalid, S. S. (2021, October 20). *YouTube trending videos: Boosting Machine Learning results using exploratory data analysis*. OUP Academic. Retrieved April 5, 2022, from <https://academic.oup.com/comjnl/advance-article-abstract/doi/10.1093/comjnl/bxab142/6404045?redirectedFrom=fulltext>
- Youtube Official Trending explanation, <https://support.google.com/youtube/answer/7239739?hl=en>
- Trending YouTube Video Statistics, https://www.kaggle.com/datasets/datasnaek/youtube-new?select=JP_category_id.json

Table 3. Summary of hierarchical regression analysis of cognitive engagement ($N = 89$ videos).

Variable	Model 1		Model 2	
	β	t	β	t
Comments	0.35	1.48	0.25	1.83
Views	0.29	1.46	0.06	0.59
Likes	0.01	0.06	-0.12	-0.95
Dislikes	0.01	0.06	-0.03	-0.01
Anger	-	-	0.18	1.43
Disgust	-	-	0.22	2.17*
Fear	-	-	0.01	0.08
Sadness	-	-	0.32	2.71**
Joy	-	-	0.50	4.59***
Surprise	-	-	0.08	0.64
Trust	-	-	1.17	8.88***
AIC	702.449		591.342	
R^2	0.08		0.78	
F for change in R^2	1.86		35.47***	
ΔR^2	0.08		0.70	

AIC: Akaike Information Criterion.

Coefficients in the table are standardized beta coefficients.

* $p < .05$; ** $p < .01$; *** $p < .001$.