

The line graph displays the percentage of respondents who believe the U.S. is a threat to their country's security, comparing the U.S. and China from 2013 to 2022. The Y-axis represents the percentage, ranging from 0% to 70% in 10% increments. The X-axis represents the year, from 2013 to 2022. The U.S. is represented by a dark blue line, and China is represented by a cyan line. Both lines show significant fluctuations over the decade.

| Year | U.S. (%) | China (%) |
|------|----------|-----------|
| 2013 | 10 | 20 |
| 2014 | 10 | 15 |
| 2015 | 55 | 60 |
| 2016 | 25 | 5 |
| 2017 | 65 | 35 |
| 2018 | 45 | 25 |
| 2019 | 5 | 35 |
| 2020 | 15 | 65 |
| 2021 | 65 | 75 |
| 2022 | 5 | 25 |

Prepared by: Rony Avivi

Table of contents:

- Introduction
- Explaining the data
- Data insights
- Conclusion

Introduction:

Background:

In the following slides I explore what characteristic of Youtube videos and channels are related to videos' popularity.

Explaining the data:

Data overview:

The dataset is taken from Kaggle and contains YouTube videos and channels metadata. The data consists of 9 direct features, and 13 more indirect features retrieved using YouTube's API.

Sample size:

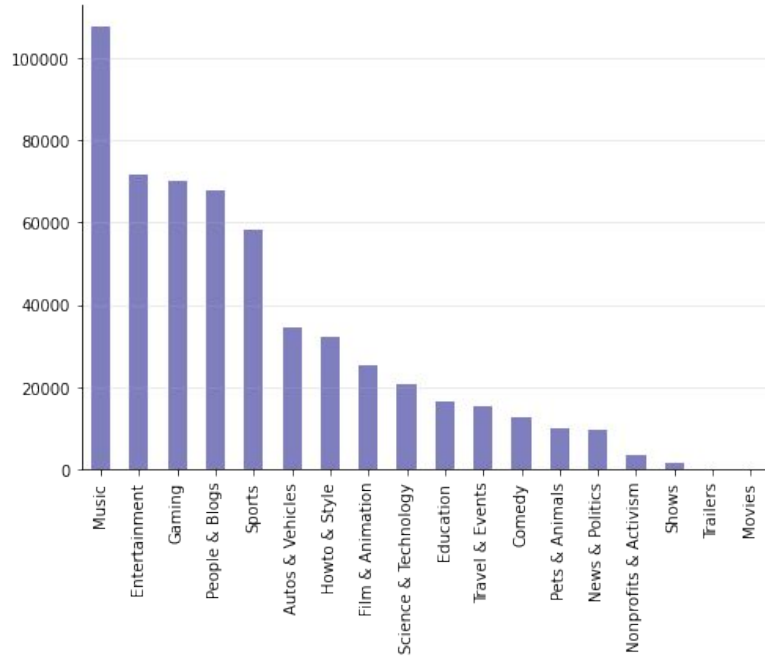
After cleaning the the data from anomalies, the data consists of 556,424 video observations which belong to 449,980 channels.

Explaining different features: Video Level Variables

| | count | mean | std | min | 25% | 50% | 75% | max |
|-------------------------------|--------|-------------|------------|-------------|-------------|-------------|-------------|-------------|
| total views / channel elapsed | 556424 | 730.31 | 6584.88 | 0 | 0.776658585 | 6.244611423 | 57.14706431 | 683532.7256 |
| channel view count | 556424 | 49446692.28 | ##### | 0 | 50621.75 | 410809.5 | 3646837.25 | 23798172642 |
| likes / subscriber | 556424 | 0.61 | 2.79 | 0 | 0.00466563 | 0.034993271 | 0.272727273 | 505 |
| views / subscribers | 556424 | 309.41 | 1587.67 | 5.38E-05 | 1.29934645 | 14.10117142 | 120.2382729 | 279803 |
| video count | 556424 | 1219.86 | 8388.58 | 0 | 17 | 70 | 338 | 415500 |
| subscriber count | 556424 | 82774.96 | 612431.04 | 0 | 38 | 411 | 5388 | 25253114 |
| dislikes / views | 556424 | 0.00 | 0.00 | 0 | 0 | 0.000185572 | 0.000554324 | 0.322981366 |
| channel elapsed time | 556424 | 67943.83 | 20632.12 | 888 | 51960 | 68400 | 85824 | 108913 |
| comments / subscriber | 556424 | 0.19 | 1.48 | 0 | 8.03E-05 | 0.003276591 | 0.045454545 | 485 |
| likes / views | 556424 | 0.01 | 0.01 | 0 | 0.001408479 | 0.003005162 | 0.006365437 | 1.418500225 |
| channel comment count | 556424 | 746.10 | 27634.98 | 0 | 0 | 3 | 24 | 3953563 |
| video view count | 556424 | 59476.98 | 1249316.56 | 1002 | 2205 | 4813 | 15296 | 500893946 |
| likes / dislikes | 556420 | 32.30 | 15.38 | 0.213044917 | 26.27240537 | 32.02631579 | 36.3 | 1835 |
| comments / views | 556424 | 0.00 | 0.01 | 0 | 0.000104147 | 0.000552091 | 0.001593202 | 1.00352963 |
| elapsed time | 556424 | 50739.67 | 19731.45 | 17520 | 34632 | 48480 | 65400 | 106609 |
| video like count | 556424 | 295.95 | 4572.45 | 0 | 6 | 17 | 60 | 1240473 |
| video dislike count | 556424 | 19.17 | 441.79 | 0 | 0 | 1 | 5 | 244280 |
| dislikes / subscriber | 556424 | 0.12 | 1.25 | 0 | 0 | 0.000400989 | 0.01369863 | 343.5 |
| views / elapsed time | 556424 | 1.54 | 31.93 | 0.009891071 | 0.045820241 | 0.106867399 | 0.347925711 | 11698.75621 |
| video comment count | 556424 | 39.48 | 563.12 | 0 | 1 | 4 | 14 | 191498 |

Explaining different features: Distribution of Video Categories

- The following graph shows the frequency of each video category by value count.
- The top five most frequent video categories are music, entertainment, gaming, people & blogs, and sports.



Explaining different features: Channel Level Variables

Channel level description of the data:

| | count | mean | std | min | 25% | 50% | 75% | max |
|------------------------------|--------|-----------|-------------|-----|-------|--------|---------|-------------|
| channel view count | 449980 | 8994588.3 | 133470264.4 | 0 | 32081 | 201405 | 1253167 | 23077225524 |
| channel elapsed time | 449980 | 67490.963 | 20647.97345 | 888 | 51648 | 67968 | 85368 | 108913 |
| channel comment count | 449980 | 157.55354 | 8925.083305 | 0 | 0 | 1 | 11 | 3953563 |
| channels with only one video | 449980 | 0.8893262 | 0.313728122 | 0 | 1 | 1 | 1 | 1 |
| subscriber count | 449980 | 21818.011 | 251298.936 | 0 | 24 | 187 | 1624 | 24430441.33 |

The table below shows the description of a new variable called 'gap', which represents the time interval between the publication of two consecutive videos on a channel.

The table includes information about the mean gap, the minimum gap, and the maximum gap:

| | count | mean | std | min | 25% | 50% | 75% | max |
|------|-------|-------|-------|-----|------|-----|-----|------|
| mean | 49801 | 318.5 | 363.2 | 0 | 74.5 | 198 | 425 | 3165 |
| min | 49801 | 258.6 | 369.6 | 0 | 20 | 106 | 346 | 3165 |
| max | 49801 | 400.6 | 410.9 | 0 | 94 | 281 | 572 | 3194 |

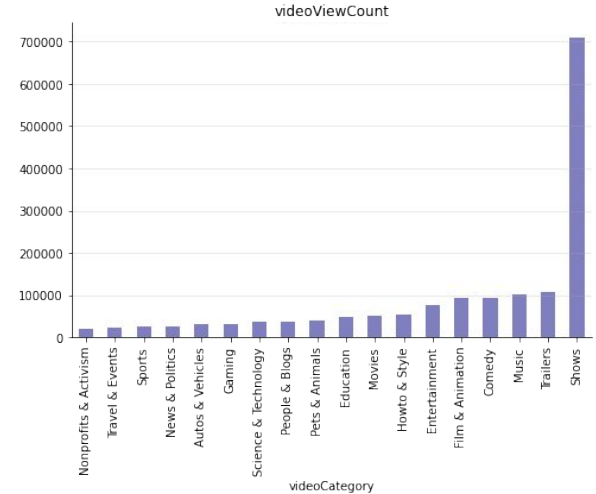
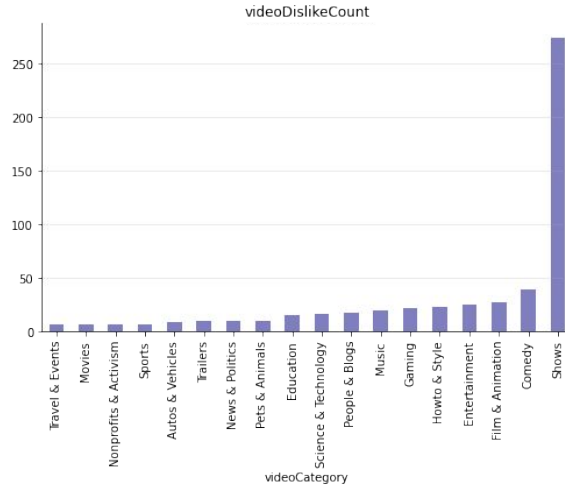
Data insights:

Let's explore which specific content or channels on YouTube lead to the higher viewer engagement.

Which categories are the most popular?

Video view count

This graph presents the distribution of video categories by view count. The category with the highest number of views is 'Shows'.

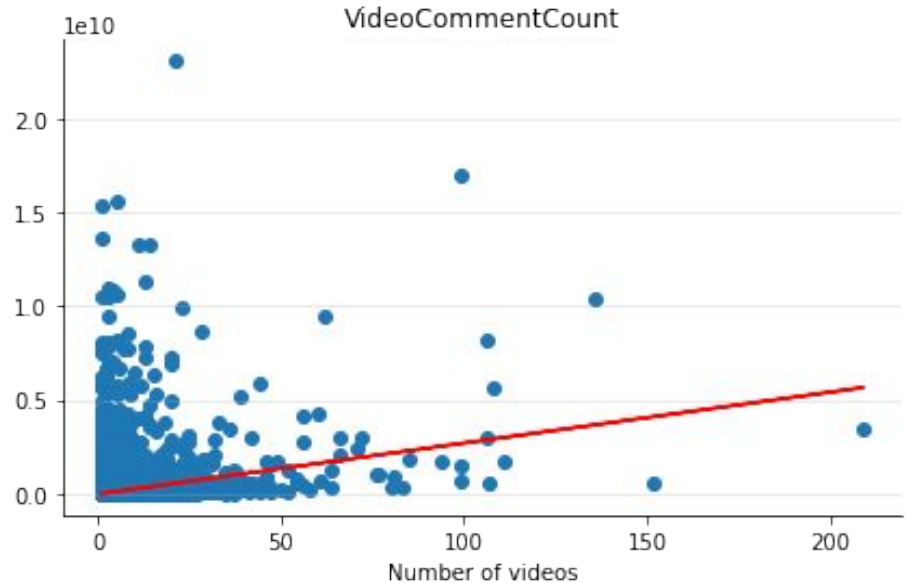


Video dislike count

This graph presents the distribution of video categories by dislike count. The 'Shows' category has the highest number of dislikes. This suggests that there may be no such thing as bad publicity.

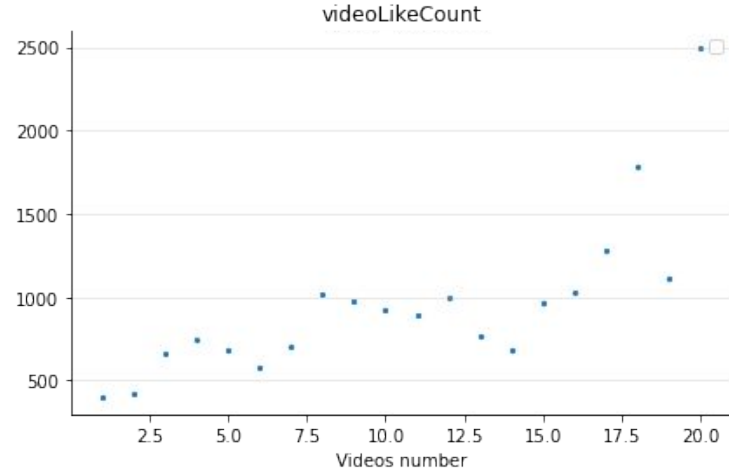
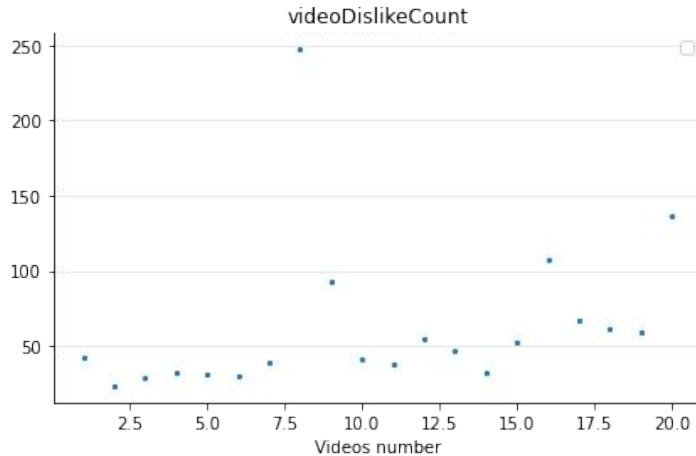
(Not surprising) Popular channels have more videos

- This graph displays the frequency of the number of videos per channel with video comment count.
- As shown, channels with more videos tend to have higher viewer engagement.



Whitin Channels, which video number is the most popular?

This case study examines channels with a total of 10-20 videos. A new feature called 'order' has been created to represent the chronological order in which each video was uploaded to the channel.



The above graphs indicate that the latest videos on a channel are more likely to engage viewers.

Conclusions:

What have we learned?

- Entertainment video categories are highly popular.
- Contrary to popular belief, a high dislike count does not necessarily correspond to a low view count.
- Channels with more videos tend to be popular.
- The latest videos on a channel tend to have higher viewer engagement.