

Business report (YouTube trending videos analysis)

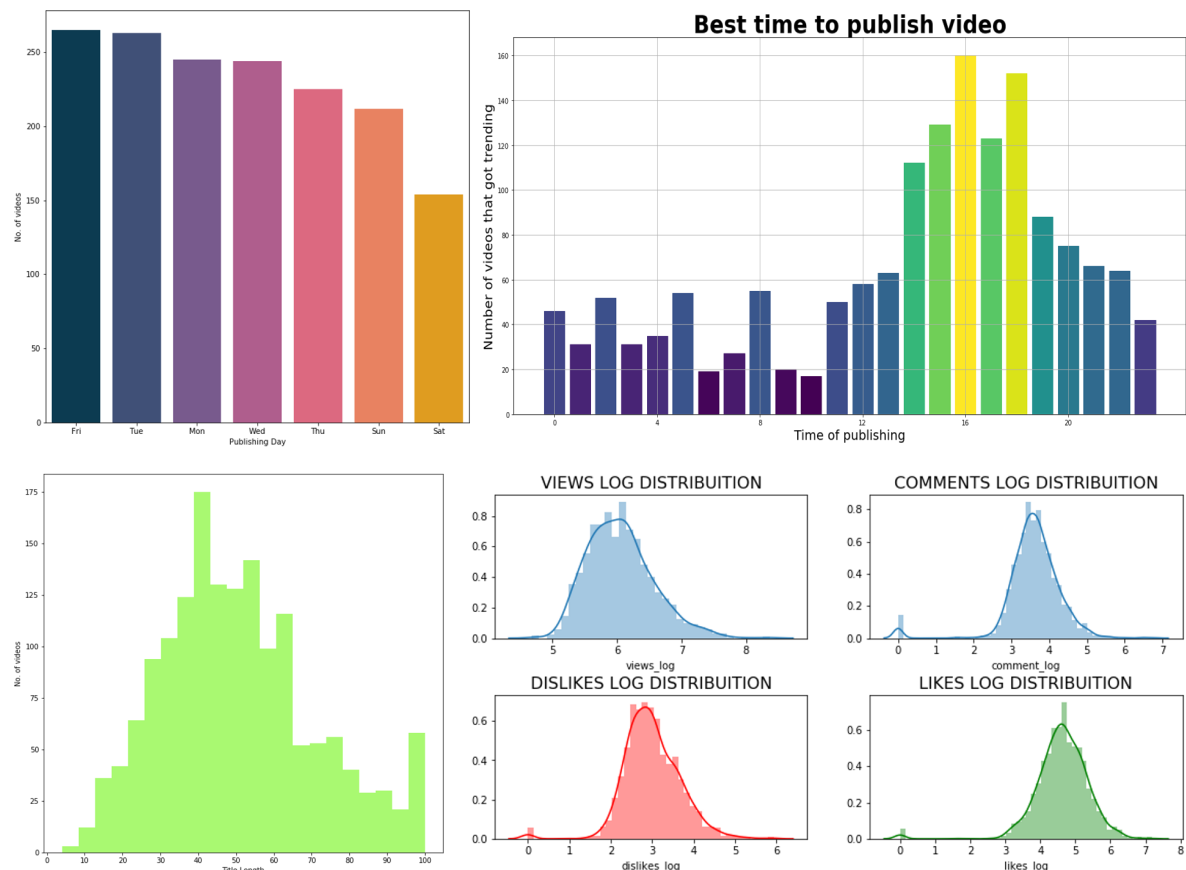
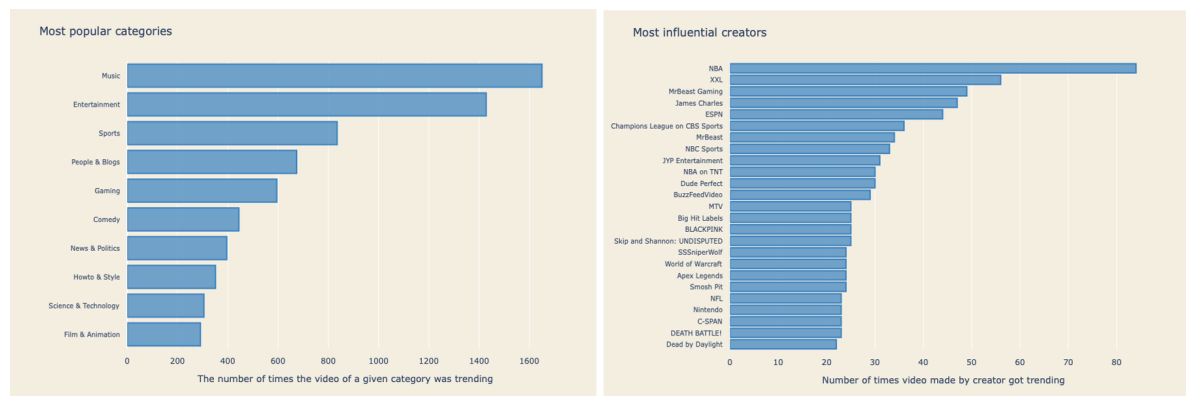
Abstract

YouTube is the largest and strongest UGC(User Generated Content) Internet video site in the world with more than 1 billion active users worldwide and the world's highest Internet user satisfaction rating. While YouTube's sophisticated channels are already booming, many new YouTube creators are facing difficulties getting off to start. In this case, we're thinking about how we can get new YouTubers to create their suitable videos and attract more views efficiently.

The first selected dataset includes data on daily trending YouTube videos in the US from August 4th to September 18th this year, with up to 200 listed trending videos per day. Data we processed include the video title, channel title, published time, tags, views, likes and dislikes, comment count, and category names. Additionally, based on the number of subscribers on September 18th this year, we select the top 50 YouTubers and choose the last 30 videos of them, dealing with their personal advantages and figures of their videos in this new dataset, including average video views, likes, dislikes, and comment count.

Analysis

Achieve the commonly trending features from Number of Videos



[illegible]

The figure consists of three vertically stacked line plots, each showing the distribution of a different engagement metric. The x-axis for all plots represents the rate (0.0 to 30.0 for like_rate, 0.0 to 17.5 for comment_rate, and 0.0 to 8.0 for dislike_rate). The y-axis represents the density or frequency of the rate.

- Top Plot: RATE DISTRIBUTION** (likes/Views). The x-axis is labeled 'like_rate' and the y-axis ranges from 0.00 to 0.10. The distribution is a smooth, bell-shaped curve peaking at approximately 0.11 around a like_rate of 5.0.
- Middle Plot: Comments/Views**. The x-axis is labeled 'comment_rate' and the y-axis ranges from 0.00 to 1.50. The distribution is a sharp, narrow peak reaching a maximum density of about 1.55 at a comment_rate of approximately 0.5.
- Bottom Plot: Dislikes/Views**. The x-axis is labeled 'dislike_rate' and the y-axis ranges from 0.0 to 1.5. The distribution is a very sharp peak reaching a maximum density of about 1.7 at a dislike_rate of approximately 0.2.

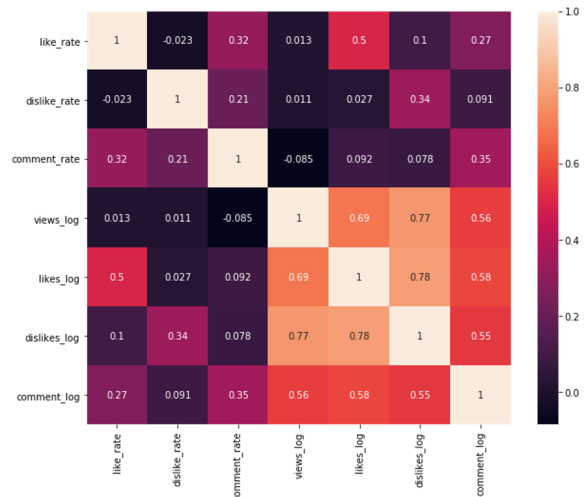
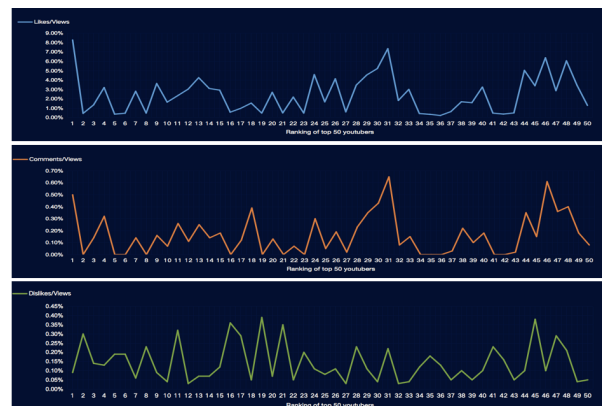


Figure 10 consists of four box plots arranged in a 2x2 grid, each showing the distribution of a different engagement metric across ten category names. The categories are: Home & Living, Entertainment, Music & TV Shows, Food & Lifestyle, Fashion & Beauty, Animals & Pets, Health & Wellness, Technology & Gadgets, Sports & Fitness, and Education & Science. The y-axis for the top row (Likes and Comments) ranges from 0 to 8, and for the bottom row (Dislikes and Retweets) it ranges from 0 to 6. Each box plot displays the median (horizontal line inside the box), the interquartile range (the box itself), and outliers (individual points below the lower whisker or above the upper whisker). The 'Entertainment' category consistently shows the highest values across all metrics, while 'Education & Science' generally shows the lowest.



On the whole, we find viewers' liking for popular videos vary from person to person, and preferred videos tend to be uniformly dispersed. However, videos that want to comment and dislike are relatively concentrated. Additionally, dislikes and views are more strongly associated, which means lots of trending videos reviews vary.

Overall, the recommendations are as follows. Firstly, choose better and suitable categories as mentioned in conclusions, with a broader range and various content. Also, try to learn from some top YouTubers in attitudes and creations. Lastly, do not forget to set proper title length and publish videos in popular time.