Analysis of Trending Video Attributes on YouTube

Introduction

As online video platforms grow in their dominance within our daily lives, the potential to spread an idea that can impact the world through these services has reached unprecedented levels. YouTube in particular stands out clearly among all online video platforms.

It is unclear how the Youtube algorithm considers a video as trending to those without insider knowledge; this paper aims to look at different factors that may encourage certain videos to be more favored on the platform with either viewers, or the algorithm itself. YouTube videos have a variety of factors that contribute to their ability to garner views as well as indicators that demonstrate viewer count and engagement. In order to gain a better understanding of what constitutes engagement, various statistics were analyzed to get a better understanding of commonalities between trending videos. The goal of this paper is to gather insight into what one can do to increase their chances of high viewership when creating, and uploading a new video. This analysis will focus on likes, dislikes, comments, categories, and tags of a video.

Data Description

A dataset, collected through the use of the Youtube API, containing a daily record of the top trending videos within the United States was used in this analysis. This data is taken from https://www.kaggle.com/datasets/datasnaek/youtube-new. It contains information on the trending date, views, likes and dislikes, comments, if comments or rating are disabled, tags, and more regarding each video that was considered trending. The table dimensions are 40,949 rows by 16 columns. The variables of interest in this paper are number of likes and dislikes, number of views, number of comments, whether or not comments were disabled, and keywords/tags used.

Results

User engagement is generally thought to correlate to engagement with commenting, liking, and disliking a video.

Relationship Between Views, Likes/Dislikes, and Ability to Comment
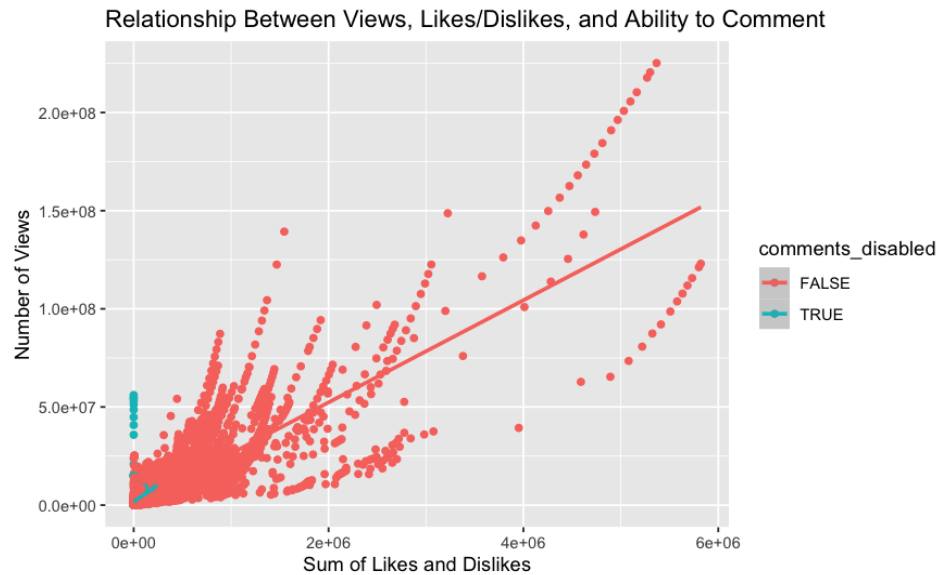
The data demonstrates that a strong correlation between the number of likes and dislikes on a video with the total number of views exists. Although high views will inevitably increase the potential number of likes/dislikes on a video, the data demonstrates that there is a strong chance that the YouTube algorithm favors videos with higher total like/dislike sums.

Additionally among the 40,949 videos within the sample, the overwhelming majority allowed for comments to be enabled. More specifically, the number of comments can be scrutinized to compare if a similar trend exists with respect to viewer engagement in this aspect of the platform.

Relationship Between #Comments and #Views
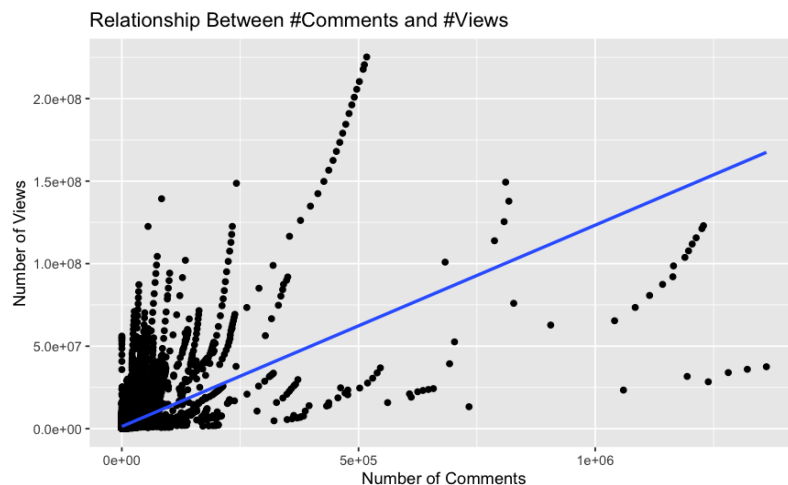
The plot demonstrates as expected in accordance with the previously presented data that the YouTube algorithm will promote videos more as engagement increases. Additionally, this plot demonstrates that the algorithm from 2017 to 2018 slightly favored comments as a more positive

indicator of a successful video to promote than it did with likes/dislikes. This is demonstrated by the slightly stronger correlation between number of comments and views, as opposed to the total number of likes/dislikes and number of views. It would be beneficial for content creators to think of videos that will stimulate discussions and encourage viewers to comment

The next variable scrutinized was the keyword/tags used for the video. Depending on the video's keywords, certain tags may have greater potential to promote a video for a viewer to watch as a recommended video or through a search bar result.
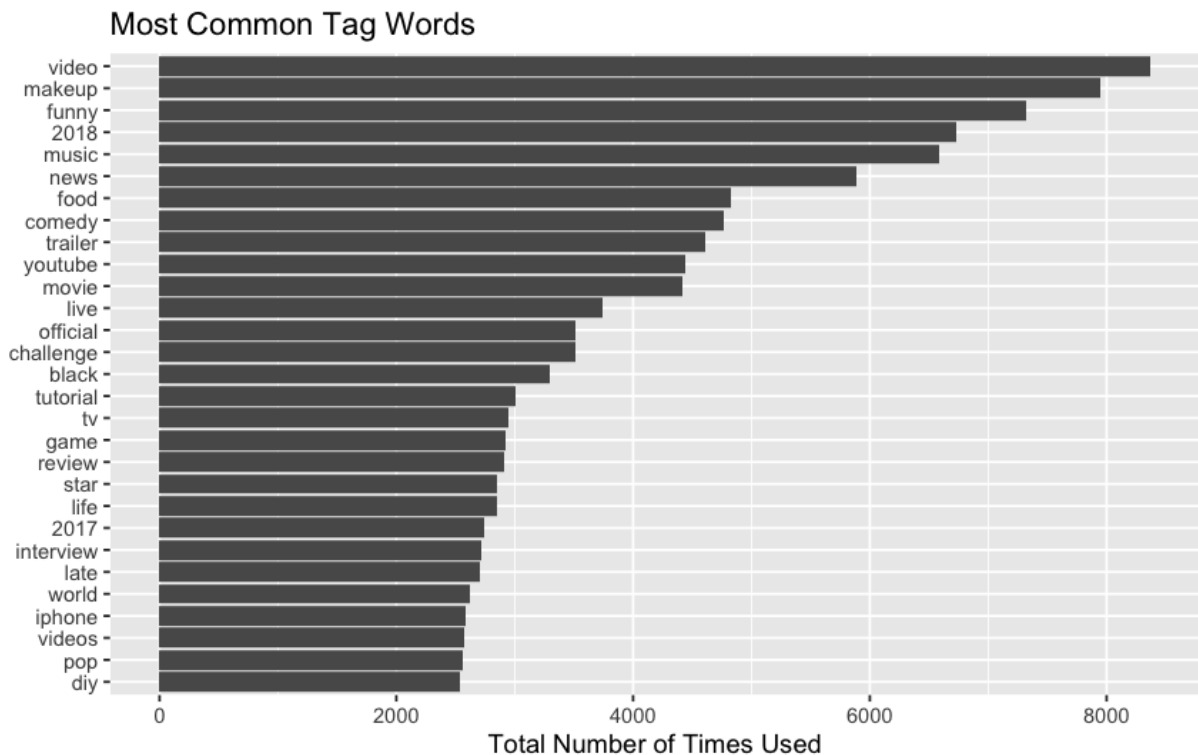
The plot demonstrates the most common tag words among trending videos within the dataset. The words depicted were thought of as individual rather than a total series; for example, the phrase "art video" was counted as "art" and "video," therefore the words depicted in the plot can be thought of as both individual and as a word commonly used in certain phrases. The most common word "video" most likely is used as previously mentioned. Notably, the following top keywords used were: makeup, funny, 2018, music, news, food, and comedy. "Makeup" demonstrates that makeup related videos were extremely popular during this time; however, the other top words are descriptive categorical words that are more general labels rather than specific, detailed descriptions of a video. This may suggest that in addition to adding tags that clearly depict your content, it would be advantageous to add more general keywords to increase the likelihood of a video being recommended to broader groups of viewers. This chart also suggests that music and entertainment related content (e.g. such as comedy, movies, etc.) tend to garner the most attention.