Team Project Proposal by Group 8

YouTube Trending Video Analysis

Team Members

Tamalika Basu
Yilei Ge
Ching-Wen(Jenny) Huang
Sungho Lee
Shulang Ning

What is your project concept?

Our team is dedicated to developing a comprehensive dashboard that analyzes the trending YouTube videos from the U.S. region. The analysis will cover various factors and attributes that potentially affect the popularity of a YouTube content creator. People like marketing analysts, business analysts, and influencers can use our analysis as references to assist their decision-making.

Who are your users?

- Marketing analysts.
- Social media activists.
- Existing YouTubers who are looking for content improvement.
- New YouTubers who are interested in competitive analysis.

What are some tasks they will do with your visualization?

As marketing analysts and YouTubers can both become potential users of our visualization, there are also several kinds of scenarios that illustrated the usage of our visualization, two of which are:

1. **Analyze if posting videos on a certain day/time of the week generates more visibility and thus more views for a video.**

   Analysts (Social Media Analysts, Marketing Analysts, Research Analysts, etc.) who are interested in marketing research can definitely gain insight from our work. Our analysis will consider a sample of data extracted from YouTube and analyze each hypothesis to check if there is a trend around different topics, such as the relationship between a certain time window and video views, the relationship between subscribers, views, likes, comments and the number of videos, etc. For analysts who are new to the field, our visualization plays a key role in providing a general understanding of the Youtube video market.

2. **Analysis of keywords that best to put in the title of the video for popularity purposes.**

   Not only analysts but also YouTube content producers will be able to gain insight through our visualization. For example, content creators can pack and classify the videos more efficiently by identifying the titles, categories, and hashtags of videos that can be stochastically popular through the visualization we provide. In addition, producers may also wonder what the most efficient video duration is. They will be able to get a lot of help in making popular videos if they

check the charts we provide, such as the most efficient duration for the video to gain popularity in their video editing process. By doing this, they can generate revenue by learning the tips for getting high view counts/engagement. With our visualizations, any creators willing to post their video and aiming for view counts can plan to release their content at the peak hours (if any) resulting in a better yield.

What is an example of an insight you hope your work will show?

1. **Hypothesis: Certain hashtags are very "powerful", which can reach more subscribers. Thus, some hashtags are strongly correlated with YouTubers' success.**

   Using hashtags properly can bring more exposure to the underlying users. In general, system algorithms (Machine Learning) can consistently predict what users want to watch by analyzing their historical "clicks". On top of that, it will suggest related videos based on users' preferences. One of the most popular ways to predict is using content categories. For instance, if someone has been watching "funny" videos on YouTube, the algorithm will learn the behavior and push more "funny" content. Therefore, videos labeled as "funny" have a higher chance of popping up as the next. Our analysis will investigate the correlation between hashtags and successful trending videos and find the keywords that are mostly used.

2. **Hypothesis: The relationship between a YouTuber's popularity (or subscribers), videos popularity (or views, likes, comments) with posting time/day trend, and amount of videos/channel is highly correlated**

   Since each YouTuber has different popularity and a different number of subscribers and videos, we want to gain insight into how high productivity leads to higher popularity. Also, since each of their videos has different views, likes, and comments, we want to learn the relationship between YouTuber popularity and the number of videos, views, likes, and comments on each of these videos. The more videos a YouTuber posts, the more views, likes, and comments will be posted by audiences, resulting in more subscribers and gaining popularity for the YouTuber. Moreover, the YouTubers will post the video on a different day and time, and the audiences will not view and like the video on the same day and time.

   This leads us to another question - when would be the best day and time to post the videos? By the youtube algorithm, if a YouTuber posts the video last, their name will be at the top of the list. Also, if the subscribers are online and want to watch the videos during this time, it will increase the possibility for the subscribers to watch them, leading to an increase in the views. Therefore, we have an assumption that if YouTubers post the video before every meal and in the evening, it will increase the number of views. Based on our hypothesis, people are likely to watch videos while eating and evening time before bed.

Where will you get your data?

- Main dataset has been found on *Kaggle*.
- We will explore more data sources if required for our analysis.