

Trending YouTube Video Statistics

Yuzhe Huang

Objective

YouTube is one of the largest video hosting websites in the world, which has profound influences on the society in all aspects. Therefore, analyzing of YouTube's dataset become significant for companies to analyze the social trends and make a prediction for future strategic business planning.

My overall objective is to provide business recommendations for investors and advertisers, including which channels have more business value and what people are concerning about.

Firstly, I will help them find the category that have relatively higher business potential. I drafted four judgment criteria, which are the number of videos in each category, average view count, variance of view count, average time interval. Secondly, we will help them find out hottest channels by four drafted judgment criteria including the appearance frequency, like to dislike Ratio, growth rate, sentiment polarity. The goal is that advertisers can choose to sign contract with these channels to make profits. In addition, they can choose these channels/YouTubers to see if they match the characteristics of their products and find potential customers. Finally, I will summarize and make predictions on what factors affect how popular a YouTube video be, which will also be valuable information for advertisers to produce viral videos for their target audience.

Dataset

- The dataset includes daily trending YouTube videos by region.
- Time interval: 2017.1-2018.5
- Data is included for 10 regions, i.e. US, GB, DE, CA, FR, JP, KR, MX, RU and IN, with up to 200 listed trending videos per day by region.
- The main variables of the dataset contains video id, trending date, video title, channel title, category id, publish time, tags, views, likes, dislikes, comment count and description.
- Source: <https://www.kaggle.com/datasnaek/youtube-new#header>

Research Questions

1. Category selection: which categories are most popular (by region)?
2. Channel selection: which channels are most popular (by region)?
3. Trending lifecycle: how long does a video keep viral (by region)?
4. Prediction: what factors affect how popular a YouTube video be (by region)?

Potential Problems

- Find an accurate model to predict the popularity of a video;
- Lack comments from viewers to make sentimental analyses for the trending videos.

Future Directions

- Dive deep into the comments to improve the analyses;

- Compare the characteristics of YouTube platform with others using the same methods to see different ecology of each online video platforms. If possible, I would write a report on the online video industry and become one of my portfolios for my future employers.