Strategies Behind the Trending Youtube Videos

Shuting Ye, Xinyang Zhang, and Yuxin Liu

University of Southern California, Los Angeles, USA
yeshutin@usc.edu
xinyangz@usc.edu
liuyuxin@usc.edu

Abstract. There are millions of users, creators and videos on Youtube. We want to discover the strategies behind these trending youtube videos. In the project, we use python to do the data processing, and d3 to draw the graph with animation and interaction also we embed the tableau page into Angular, javaScript framework, and bootstrap, front-end framework, to demonstrate some basic attributes.We built the responsive website to demonstrate the best time for publishing, the average time for becoming trending, different country's preferences for categories, the performance for the same categories in different countries and some detail statistics includes like, dislike, views, comment count infographics. The above information would inspire the user to customize their video in a more efficient way.

Keywords: D3 chart · Transition · Responsive · Tableau · Youtube.

1 Introduction

More than 1,300,000,000 people are watching Youtube video in the worldwide, 300 hours of video uploaded to Youtube every minute.[1] There are over 23 million YouTube Channels and part of them are full-time Youtube creators and part of is just uploading it for fun. From this data set, we want to find the strategy behind these trending videos, like what is the most common used tags in the world, which time is the best time for publishing the video, what is the temperature for different categories among 9 countries and how long would it become trending after publishing? Based on the data, we want to provide the information about the categories in different countries, the time for publishing and the tags for the video to those who want to make their videos become trending.

2 Data

2.1 Collection

The data set comes from in Kaggle [2] which presents the 9 countries' trending videos statistics, including France, Germany, United Kingdom, India, Japan,

Korea, Mexico, United States, Canada which belongs to Europe, Asia and South America. The dataset includes data gathered from videos on YouTube that are contained within the trending category each day. There are two kinds of data files, one includes comments and one includes video statistics. They are linked by the unique video_id field.

2.2 Processing

Aggregation and collation of data was a complex process considering there are sheer volume of the data and the amount of processing required. Some basic data processing was performed on the data like Python NumPy and Pandas. The description and tags were split and group by each country. Also, we filtered out lots of missing value entries and eliminated duplicate rows.

3 Approach

3.1 Design

The design of charts focuses on the upper half of Cairo's visualization wheel. It has been built based on the design principles of Cairo's wheel. It shows the different time, category, tag the trending videos common have.

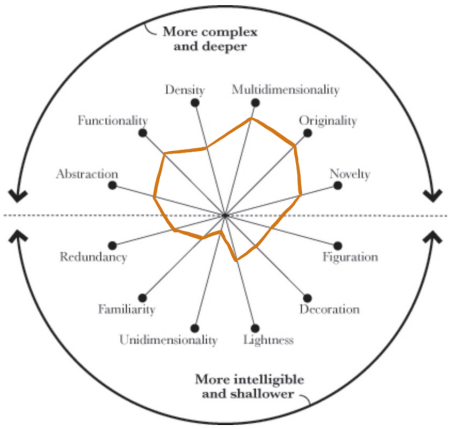

Fig. 1. Cairo's Wheel

This helps the user to get a good overview of the overall data so that they can click to choose and know more detailed information, making it multi-dimensional. Also, the visualization is highly functional instead of being decorative. The graphs are also not redundant and display novel aspects.

3.2 Word Cloud and Bar Chart

The word cloud graph would present the most commonly used tags that trending video have. Users could look into the bar chart to find the different categories and they can click on different bars to show the word cloud for the most common tags in this category. And the shape of the word cloud would show in the youtube logo style.

3.3 Bubble Chart and Tree Map

The treemap was built by Tableau tool and it demonstrates the basic attributes includes likes, dislikes, views, comments about the videos and also the bubble chart draw by D3 used the bubble radius to present the relationships among the likes, dislikes and views. We want to use the treemap to show the statistics about the basic data and use bubble chart to show the distribution of data.

3.4 World Map and Donuts Chart

The world map shows the temperature of different categories among 9 countries which means show the popularity degree. Firstly, users could choose from top 5 categories button to see the color change on the map. And when users want to discover the certain country's top 5 likes category, they can click on the country path on the world map which would present a more detailed and large donuts chart to show the exact top 5 categories for that countries.

3.5 Lollipop chart and Line chart

The lollipop chart could present the smooth transition to present the different publishing time on weekdays and in hours. Users can choose from the drop-down menu to customize their graphs. And also the line chart could present which weekday accelerate the trending speed of youtube videos most among 9 countries. When you click on each button for different countries, and y-axis could transit in the smooth way towards different scale.

3.6 Bar Chart

Bar chart could exhibit the speed of a video from publishing to trending and the graph could let the user choose any category video and any countries they are interested in.

4 System

The website can be used in most browsers(Safari, Chrome) and can be rendered in all Operating Systems like Mac, Windows, Linux, iOS, and Android.

5 Related Work

Edward L. Platt, Rahul Bhargava, Ethan Zuckerman's paper show their work about the geographic popularity of videos by incorporating trending data and extending their analysis from video-nation affiliations to nation-nation co-affiliations[3]. Jonathan Scott Morgan, Iman Barjasteh, Cliff Lampe, Hayder Radha used information theoretical measures based on entropy to examine how time-series distributions of common measures of popularity in videos from YouTube's "Trending videos" and "Most recent" video feeds relate to the theoretical concept of attention[4].

6 Conclusion

The main object of our project is to visualize the information of the existing trending video's pattern and feature. The visualization was made so that almost every user could find the information easily. The user could easily get the most frequent tag word in a different category, and based on different 9 countries, user could choose the category to show the temperature in different countries. Also, within a week, creators could choose from Tuesday to Friday, and within 24 hours a day, creators could choose from 2:00 p.m to 5:00 p.m to publish their video which would become trending in the shortest period. The user also can look into the like, dislike, views and comments' distribution in the different countries and categories.

7 Acknowledgment

We would like to show our appreciation to Professor Luciano Nocera, thank for helping us to find the dataset and provide the guidance for the content and graphs.

References

1. 37 mind blowing youtube facts, figures and statistics – 2019. `https://merchdope.com/youtube-stats/`, accessed September 29, 2019
2. Trending youtube video statistics. `https://www.kaggle.com/datasnaek/youtube-new`, accessed June, 2019
3. Karayev, S., Trentacoste, M., Han, H., Agarwala, A., Darrell, T., Hertzmann, A., Winnemoeller, H.: Recognizing image style. arXiv preprint arXiv:1311.3715 (2013)
4. Morgan, J.S., Barjasteh, I., Lampe, C., Radha, H.: The entropy of attention and popularity in youtube videos. arXiv preprint arXiv:1412.1185 (2014)