The Significance of Trending Videos Through the Analysis of Variables

Stephanie DeMaria, Maegan Trygstad, and Alexis Wright

Northern Illinois University

Correspondence concerning this article should be addressed to Stephanie DeMaria, Maegan Trygstad, and Alexis Wright, Northern Illinois University, DeKalb, IL 60115.

Contact: sdemaria@niu.edu, mtrygstad@niu.edu, awright@niu.edu

**What is a Trending Video?**

In today’s society the creation and consumption of media is expedited through the use of social media sites and “sharing” based platforms. For this project we will be looking at YouTube, the largest video sharing platform to date. YouTube, created in 2005 by three former PayPal employees, is a platform where users can upload, view, rate, share, and comment on videos. The videos on this platform can range from video blogs (vlogs), movie trailers, music videos, short films, live streams, or documentaries. The reason we chose YouTube as a significant dataset was because of the amount of content that is being shared and consumed on the site. According to Alexa Internet, there are more than 400 hours of content being uploaded every minute and one billion hours of content being consumed every day. With this amount of content being shared, there is plenty of significant information that can be gathered from YouTube’s data.

Our data set specifically focuses on trending YouTube videos between November of 2017 until June of 2018. Per YouTube’s own Help site, “trending helps views wee what’s happening on YouTube and in the world.” YouTube has two distinct categories in trending, the video type and the trending signals. Video types that can gain trending status are appealing to a wide range of viewers, are not misleading, capture the breadth of what is happening on YouTube and in the world, and are surprising or novel. Trending signals that go hand in hand with video types include, but are not limited to, view count, rate of growth in views, where the views are coming from, and the age of the video. The combination of the types and signals is updated every 15 minutes to have a continuous showcase of what is popular on YouTube at that time.

The main goal of this paper and the analysis of this data set is to identify the different aspects of what makes a trending YouTube video and to compare the different types of trending videos. From our data set we are able to compare what country the video originated from, the view to like/dislike ratio, and [something]. The different variables that were included in the data set are as follows: Country, Video ID, Trending Date, Title, Channel Title, Category, Publish Date, Tags, Likes, Dislikes, Comment Count, Thumbnail Link, Comments Disabled, Ratings Disabled, Video Removed, and a Video Description. Although not every variable will be relevant to data visualization, every category is relevant to the video’s trending, so the variable was left in the data set. The relevant variables that were used for the data visualizations will be defined and their relevance explained.

**Country**

Within our data set we had trending video data pulled from ten countries, each one of this countries’ trending videos were mined in the same way, with an exception of India. According to YouTube’s Help site in India, trending displays the same list of trending videos for each of the Indic languages. We included ‘Country’ as a data variable because it is important to compare the differences in trending video between countries, such as a difference in views in different countries for the same trending video, or the difference in trending lists at any given time.

**Trending Date/Publish Date**

The publish is the actual date that it was uploaded to YouTube. The trending date is simply the day the video made it onto the trending list, it should be noted that this does not have to match up with the publish date, although it typically does. There is a phenomenon known as the ’24-Hour Record,’ for videos that gain a significant amount of views within a 24-hour period (as compared to the life of the video). An example of this is Ariana Grande’s music video for her hit song ‘Thank U, Next’ which garnered 16 million views in just under 22 hours. Long awaited music videos will typically appear on the trending page because it falls under the trending guidelines that were listed above. Other videos that fall into this category are new released movie trailers, highlight reels from recent sporting events, or uploads from famous channels. Most of these types of videos hit the trending status within a few days, the one outlier being ‘viral videos’ which are rapidly shared, this video can be from years ago but if it reached ‘viral video’ status it may gain the coverage it needs to become a trending video.

**Title/Channel Title**

The title and channel title are simple what the video was called and what the name of the uploading channel is. Although this data is not as significant in our data set, we thought that these variables could be relevant for later studies. A potential study that could be done would be to identify what videos are being viewed because of quality content/genuine interest, or if they are simply being viewed because of the channel title associated with it. This type of study could affect how trending videos are chosen, if views become a biased way of choosing trends. For our project we used the title and channel title as a way to easily sort through the different videos.

**Categories/Tags**

The category for a YouTube video is chosen by the creator, advertisers, or channel managers so that the content is identified in the way they see fit. Typically, videos are categorized in broader categories to account for redundancies or over-specification, an example of a broad category would be Food or Gaming. For the purpose of our project we used categories as a way to further separate videos and used this separation to ask questions such as, “Do certain categories get more views that others?” or “Are certain categories more disliked than others?” These types of questions are important because it can tell content creators which categories are more relevant at a given time. Similarly, tags can also help identify videos in a more detailed way, instead of Food, you could get the tags ‘Baking’ and ‘Cupcakes’ to further narrow the scope of the video. Tags help users find specific videos by generating relevant keywords that are associated with the type of video and its content.

**Likes/Dislikes**

One of the most well-known aspects of YouTube and its content is the use of a Like/Dislike system, both choices are simply identified by a thumbs-up or thumbs-down symbol. As mentioned above likes and dislikes can play an important role in what video gets chosen as a trending video. For this project we discussed how likes and dislikes correlated to different categories and [something]. One of the limits of the use of likes and dislikes is that they may not always be truthful to the actual opinions of those clicking the thumbs-up or thumbs-down buttons. There have been groups of people that target specific channels or artist and dislike their video because of a large issue, not because they dislike the video. False likes and dislikes that this can affect which videos are trending which is why they are not the only variable taken into consideration.

**Comments/Ratings Disabled**

These categories simply tell which video creators turned off the rating system or the comment section for their video. These selections can be made for a variety of reasons, such as sensitive topics, topics that don’t require discussion, or topics that may dissolve into unnecessary fighting. For this project we looked at which categories had more comments or ratings disabled to identify some of the factors in disabling these variables.

**The Relationship of Pertinent Variables**

As evident by the section above, we did not use every variable in the data set to create visualizations, instead we chose the data that was most relevant to our goal. The combination of the variables we chose come together to show us the types of videos that become trending through the use of different measurements like views and likes/dislikes, what categories are most prevalent, and how these variables are different from country to country. By using these variables together, we were able to gather data that could be used by content creators to better promote their videos and achieve trending status.

**Visualizations**

In this next section we will give an example of the different visualizations that were created to convey important information. The visualizations are grouped together by their relevance to each other to tell a story. Each visualization will be followed by a brief description of why we chose the type of chart and what the findings are from each story.

**Story 1**

**Visualization 1.**

**Visualization 2.**

**Visualization 3.**

**Story 2**

**Visualization 4.**

**Visualization 5.**

**Visualization 6.**

**Story 3**

**Visualization 7.**

**Visualization 8.**

**Visualization 9.**

**Conclusion**

From the visualizations shown above we can see

References

Trending on YouTube - YouTube Help. (n.d.). Retrieved from https://support.google.com/youtube/answer/7239739?hl=en

“Youtube.com Traffic, Demographics and Competitors”. *www.alexa.com.* Retrieved 2 December 2018