



# **Analyzing YouTube WorldWide**

## **Team 15 Capstone Project**

### **August 10, 2022**

Final Project for Correlation One  
Data Analytics  
Cohort 2

Team 15:  
Matt Hendry, Ponisha Pokharel, Theodore Gebre, Jimmy Hernandez, Amy Hart

## Introduction

With the explosive growth of content and viewership since its humble beginnings 15 years ago YouTube has become the fastest growing website on planet Earth. It has grown from a few kids putting together home videos to a top tier entertainment entity owned by Google reporting 28.8 billion in advertising revenue in 2021. Our team chose to explore the mechanics of going viral on Youtube and with a focus on 5 countries, one for each of our team members to analyze.

## Business Questions:

How do Likes, Dislikes and Comments affect view counts on YouTube?”

What are the top 5 categories for each country?

What are the mechanics of going viral?

## Target Audience:

Our Target Audience is advertisers and content creators. Keeping track of what's trending and what current categories enjoy the largest view counts reveals the leading edge of interest in YouTube content. The creators and their channels become safe bets for a larger audience share. For our purposes we cleaned our data set with this in mind. Youtube in general is a sure bet for advertising just by the sheer size of its audience. Its very niche is viral, ad space is short and very specific to the YouTube Audience. Trending times and viewer engagement statistics can help to direct placement of ads for specific products. Creative content feeds advertising, and advertising feeds creative content. It truly is a win win situation!

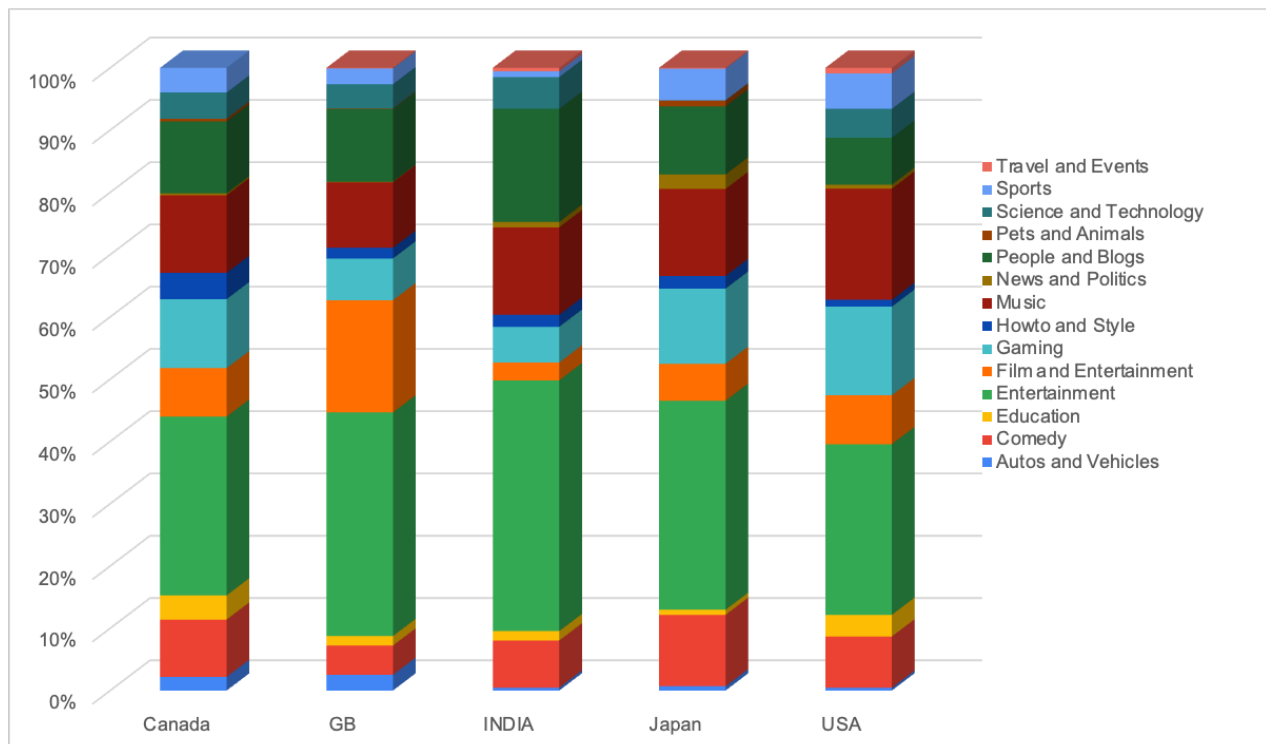


## The Data

We found our data source on Kaggle. The data set was scraped from a live data base that is updated everyday. We knew we had to draw down our focus as the data set was too large for us to use effectively so we settled on choosing 5 countries and a common parameter of all videos uploaded in the month of December 2021 in those countries. We downloaded our sets and got to work cleaning our data.

<https://www.kaggle.com/datasets/rsrishav/youtube-trending-video-dataset>

It is Ironic that the most popular videos on YouTube are on a channel known as Mr. Beast. Our data set is one heck of a Beast. The amount of data is so big that no one wanted to download it. So we split it up and found the categories/columns we wanted. We each took our country's data and cleaned and sorted and came up with our project dataset which we loaded into excel at first to wrangle. Our data columns were sorted and we focused on View Counts, Likes, Dislikes, Comments, Trending Dates and Publishing Dates for our project.



Each Team Member cleaned their own data set:

Matt Hendry - Canada

Ponisha Pokharel - India

Ted Gebre - Great Britain

Jimmy Hernandez - USA

Amy Hart - Japan

We then recombined the data from all five countries to one data set.

We decided to narrow our focus to the top five categories, which were almost the same for each country, The ranking is a little different in different categories, and we combined some categories.

Entertainment is a huge category, and has by far the most view counts. But Entertainment is not as specific as the other categories, and things from other categories end up in the Entertainment category, like Film, Animation, Comedy, Movies, etc. Whereas Music is just music, and Gaming is just gaming. We took this into consideration and decided to get more specific. To gain a better perspective of the category content, We narrowed our focus to the top 5 videos in each of the top 5 categories. Our data set went from 6581 rows to 125! We were able to see marked differences and great similarities in views and trends per country.

Our dataset included Published Date and Trending Date per title, so we developed a formula to use that information to track the trending speed. Again we averaged it per category and country to use in our project as a sort of key performance indicator. There were some very interesting findings. 4 out of 5 countries were dominated by Mr. Beast! Our analysis was not necessarily focused on the titles, but leans toward the process of identifying the trends of countries as they related to the category, and viewer engagement numbers.

## The Dashboard

We took our dataset to Tableau as well as the smaller version with the top fives.

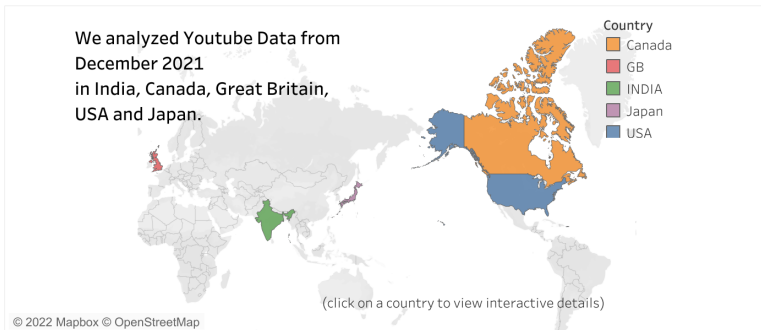
<https://public.tableau.com/app/profile/amy.hart4142/viz/>

### Analyzing YouTube Worldwide

Team 15: Matt Hendry, Ted Gebre, Ponisha Pokharel  
Jimmy Hernandez, Amy Hart

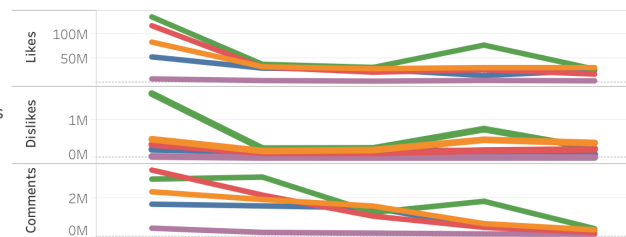


#### Where In The World?



Likes dislikes and comments tend to follow view counts, but there are unique and interesting finds between categories and countries.

#### Likes Dislikes and Comments



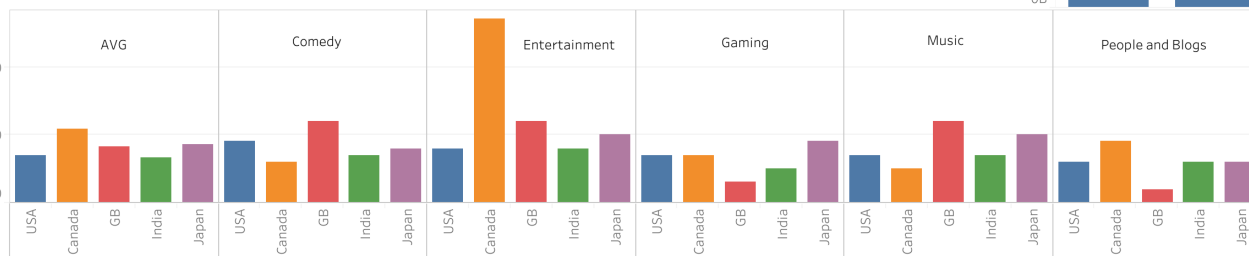
A focus on the top five categories in each Country based on View Counts.

#### Top 5 YouTube Categories December 2021



This chart shows the average length of time in days in each country / category from Publish to Trending. In terms of the mechanics of becoming viral we can see the speed to trending in each category/country and an average of all categories by country shows us that the averaged viral speed in India is the fastest of this group.

#### Trending Average by Days (Less is More!)



[Team15AnalyzingYoutubeWorldwide/Dashboard1](#)

With the big data set we were able to generate an interactive map with each country's statistics. This way we can easily see how viewership is spread across nations, and categories.

In the lower chart, we used the smaller data set to drill down on trending averages to show how the speed of trending averages between countries.

Next we factored in population and did some calculations using all the data to find relationships between country sizes and viewership numbers. In some cases it was obvious that size mattered, but in other cases looking at specific categories gave us unique perspective as to whom was watching what. For example, although Canada and Great Britain are smaller population wise, their view counts were very large in Entertainment and Comedy respectively.

Rank	Country	Continent	Population	Source
2	India	Asia	1,407,563,842	<a href="https://population.un.org/wpp/">https://population.un.org/wpp/</a>
3	United States	America	332,947,467	<a href="https://www.census.gov/popclock/">https://www.census.gov/popclock/</a>
11	Japan	Asia	125,502,000	<a href="https://www.stat.go.jp/english/data/jinsui/2.html">https://www.stat.go.jp/english/data/jinsui/2.html</a>
21	United Kingdom	Europe	67,081,234	<a href="https://www.ons.gov.uk/">https://www.ons.gov.uk/</a>
37	Canada	America	38,857,997	<a href="https://www.census.gov/popclock/">https://www.census.gov/popclock/</a>

We answered our business questions in the following ways:

Likes, Dislikes, and Comments definitely help content to trend. They are symbiotic. One feeds the other and vice versa, just like a biological virus, needs a host, a country, a creator, a channel, an audience. Viewer engagement affects Categories in different countries in different ways for instance, viewers in Japan don't hit the dislike button like the viewers in the USA. It can also be said that the viewers in Japan tended to comment more on People and Blogs relative to the view counts of more popular categories. The nature of People and Blogs is one of dialogue, and engagement.

The Top 5 categories with the most view counts per country for the Month of December are:

1. Entertainment
2. Gaming
3. Music
4. Comedy
5. People and Blogs

The rankings of these categories is differs between countries.

The Mechanics of Going Viral:

Likes, Dislikes and Comments feed the algorithm. The time between the date of upload or publishing date and the trending date are considered the speed of the spread of viewer engagement. It is interesting to note that the most popular categories do not always trend the fastest. The size of the population, as well as climate may have some bearing on the numbers. In countries with more precipitation, namely Canada with its long winters and Great Britain with its famous rainy weather, climate may be a factor in why these smaller populations have larger view counts.

Overall it was a fun project. We learned so much about YouTube and the countries we studied. We wish we had time analyze all the countries that enjoy YouTube in our research.

## Sources

<https://www.kaggle.com/datasets/rsrishav/youtube-trending-video-dataset>

<https://www.kaggle.com/datasets/datasnaek/youtube-new>

<https://www.statista.com/statistics/277758/most-popular-youtube-channels-ranked-by-subscribers/>

<https://www.businessofapps.com/data/youtube-statistics/>

<https://www.worldatlas.com/articles/which-countries-watch-the-most-youtube.html#:~:text=Top%2010%20Countries%20with%20the%20M>

<https://www.tubics.com/blog/what-counts-as-a-view-on-youtube>