# Team 022 Project Description

**Describe what data is stored in the database. (Where is the data from, and what attributes and information would be stored?)**

We will be using the TA-Proposed database “Youtube Trending Video Dataset” Along with the data we calculated with the data in this dataset and the data we retrieved from the Youtube API.

This dataset includes data on daily trending YouTube videos in multiple different regions such as India, USA, Great Britain, Germany, Canada, France, Russia, Brazil, Mexico, South Korea, and Japan. with up to 200 listed trending videos per day.

Specifically, the data includes the videoId,video title, channel title,channelId,categoryId ,publish time, tags, views, likes and dislikes, description, comment count,trending time, comments_disabled, rating_disabled, descriptionvideo_id, thumbnail links,region.

We will retrieve/calculate things such as: like/dislike differences, top comments,comments with the most likes, comments with the most reply, comments with replies from video providers, average view counts, average view time, regional view difference with given tag,url,duration,region restriction, and so on.

**What are the basic functions of your web application? (What can users of this website do? Which simple and complex features are there?)**

Choose a country to get data from, find out if they want supported/disapproved subjects from that respective company and what the top liked comments are (will help show culture). There will also be visualizations in the form of pie charts and bar graphs. I think the complex functions will be those data analytics while the simple part is querying the right data from the user pressing a button for that specific data (might be in the form of a check box.)

Our possibilities are below:

1. We will have options to choose from and make a bar graph
   Ex: country and their likes/dislikes top subjects plot
2. Comparing two or more specific videos and shown using a table Ex: Drop down list

3. Pie chart of regions and their top subjects (using counts of words in trending video descriptions and titles
4. Average watchtime line graph for trending videos of certain regions
5. Focus on likes/dislikes and correlations with comments disabled
6. Seasonal video analysis like during certain month what subjects appear the most (ex: seeing if christmas videos trend most during December in different regions)

**What would be a good creative component (function) that can improve the functionality of your application? (What is something cool that you want to include? How are you planning to achieve it?)**

A good creative component for functionality is maybe a way to cache the user's region based on the inputted data he/she wanted to look at. This will then generate graphs and pie charts of similar data from that region to kind of give personalized ads to the user based on the earlier queried data. Something cool specifically for the frontend we want to include is like a pie chart animation that is part of the homepage, showcasing top subjects that trend for different regions.

**Project Title**

***Youtube Trend or Youtube Manipulation?***

**Project Summary:  It should be a 1-2 paragraph description of what your project is.**

The project will revolve around analyzing what is being fed as content to different countries. This analysis will be about what content and reactions certain topics have from certain countries. We can identify whether Youtube classifies certain topics with certain countries and how that might in turn grow that culture in that country. Our project will allow a user to discern and identify content that is being fed to them and other countries through visualizations and statistics.

We will specifically use Likes, Dislikes, Video Descriptions, Video Descriptions, Tags, Views, and much more to discern public reaction (acceptance and angry pushback). This will make a database that will store information that can figure out what each country is classified as in terms of its subject approval and disapproval. These trends will be our focus.

**Description of an application of your choice. State as clearly as possible what you want to do. What problem do you want to solve, etc.?**

We will use the trending video data to discover the differences between trending videos in each country (labeled in the region tag), and analyze what topics seem to be most relevant in their respective countries as well as the way certain topics affect the thinking patterns of citizens in those countries (we can check the comments on the video). Basically we want to identify the ways that Youtube's Trending feature may be manipulating the public's thinking in various countries.

**Usefulness. Explain as clearly as possible why your chosen application is useful. Make sure to answer the following questions: Are there any similar websites/applications out there? If so, what are they, and how is yours different?**

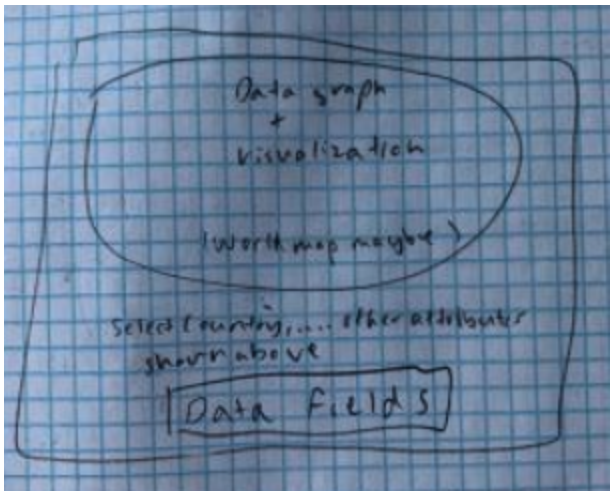
We can identify if Youtube is promoting certain agendas/ideologies in certain countries, which would be useful as an informative tool for people in the public who wish to know if that is the case. While there are Youtube trending analysis tools, there are no well known tools that specifically try to identify agendas/ideologies that Youtube may be forcing upon the public. Compared to these trending analysis tools, which are typically used to get basic information such as like to dislike ratio, our application would provide far more information about multiple videos that follow a certain topic trend.

**Realness. Describe what your data is and where you will get it.**

We will get the data from Kaggle's Youtube Trending Videos Dataset, which "includes several months of data on daily trending YouTube videos" from various regions and up to 200 trending videos in any given day. It possesses video title, channel title, the time the video was published, any tags, the view count, the like and dislike counts, the description, and number of comments. We can also utilize the video titles to gather other data outside of the Kaggle data, such as video length, content of comments, etc.

**Description of the functionality that your website offers.**

- **A low-fidelity UI mockup**: What do you imagine your final application's interface might look like? A PowerPoint slide or a pencil sketch on a piece of paper works!




- **Project work distribution**: Who would be responsible for each of the tasks or subtasks?
  List of the person responsible for which exact functionalities in section 6. Explain how backend systems will be distributed across members. Be as specific as possible as this could be part of the final peer evaluation metrics.
- Functionalities
  - We will have options to choose from and make a bar graphEx: country and their likes/dislikes top subjects plot
  - Comparing two or more specific videos and shown using a table Ex: Drop down list
  - Pie chart of regions and their top subjects\
  - Web scraping for comments data
  - Find most common youtuberfor regions or average watchtime line graph
  - focus on likes/dislikes
  - correlated with comments disabled
  - average length of comments
  - per region
  - we should aggregate watch time by ranges

| **Name** | **Task** |
|---|---|
| Nihal | Front End Data Insertion and Retrieval with SQL queries |
| Derek | Sort Data based on prioritization |

| | |
|---|---|
| Akash | Connect Tables for Queries and work ond front end data visualization and connectivity |
| Kevin | Clean Data to Remove Certain Values based on Prioritization |