US vs GB

1) The purpose of the project

This project aims to analyze a decade worth of YouTube data on USA and GB. Our team will look into similarities, and differences, on two countries' viewing preferences that may provide valuable insights to those who plans to expand their contents on either country.

2) A background on the data used

- Graphs: Chart.js, for all graphs
- The dataset was obtained from Kaggle:
  https://www.kaggle.com/datasnaek/youtube-new
    - It contains the following fields that will be utilized for this project:
        - Trending date in date format
        - Channel as string
        - Category id as integer
        - Likes, dislikes and comment counts as integers

3) Hypotheses and expectations of the final visualizations and conclusions.

- We selected US and GB due to the similarity in data volume and were hoping we had consistent data from 2007 -2017.
    - After the further review of the data the years available for each country were not consistent and had to change the approach on the line graph to show more cohesive data points.
- The hypotheses were that the US was going to oversee the You Tube users in all categories.
- Once digging into the data, we decided the following visualizations would best show You Tube categories trends in US vs GB for You Tuber's
- Scatter Plot- demonstrated the likes and dislikes per category. The goal of this visualization is to present the "likeability" of each category to display popularity.
- Bar Graph- demonstrated average views per category and the expectation was to focus on the most viewed categories to assist You Tubers on what areas are best trending amongst viewers to support future video uploads.
- Line Graph- demonstrated average views per year and the expectation was to show which country was more dominant to in return assist You-Tubers with crossovers between countries.

- Doughnut Graph- This graph compares the top ten most liked YouTube categories based on the number of likes cumulatively received. The purpose of the graph was to provide a summary between both countries and view similarity totals for the top 10 categories

4) A technical description of the application with a paragraph dedicated to each viz. This should also include libraries used, etc.

Description for: Homepage
The application is
The front end of this application uses HTML and CSS. A starter template was utilized. Customization on the template was done to render a professional looking frontend while reflecting the team's vision and goal.
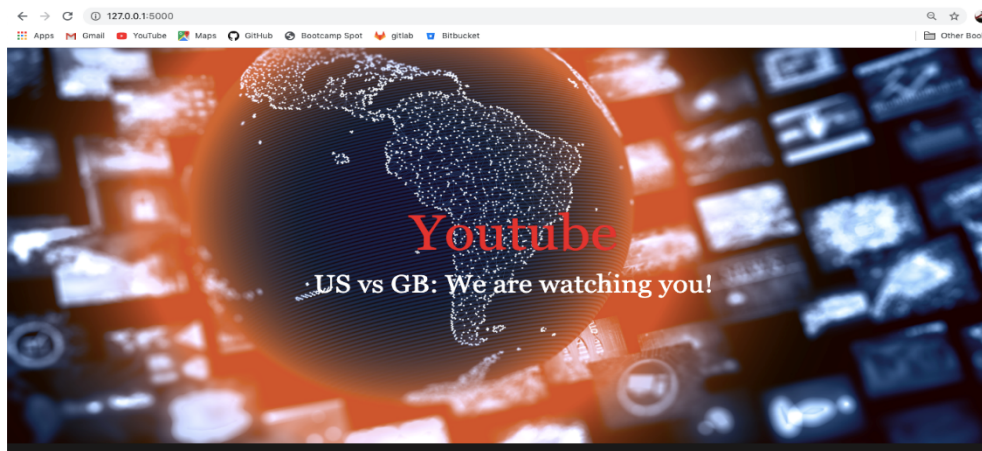
At the backend of this application are Python and Javascript. Python was used on the Extraction, Transfer and Loading process of the data to the database. It was also used to create the flask app that run this application. Javascript was utilized on rendering the visualization with the aid of the Chart JS library.
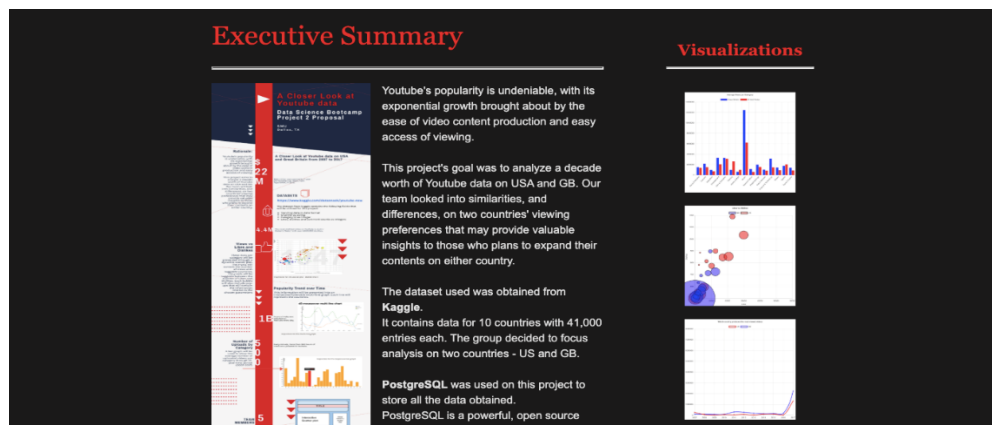
The dataset used was obtained from Kaggle.
It contains data for 10 countries with 41,000 entries each. The group decided to focus analysis on two countries - US and GB.

PostgreSQL was used on this project to store all the data obtained.
PostgreSQL is a powerful, open source object-relational database system that has earned it a strong reputation for reliability, feature robustness, and performance.

Description for View per graph:

The application is

The front end of this application uses HTML and CSS. A starter template was utilized. Customization on the template was done to render a professional looking frontend while reflecting the team's vision and goal.
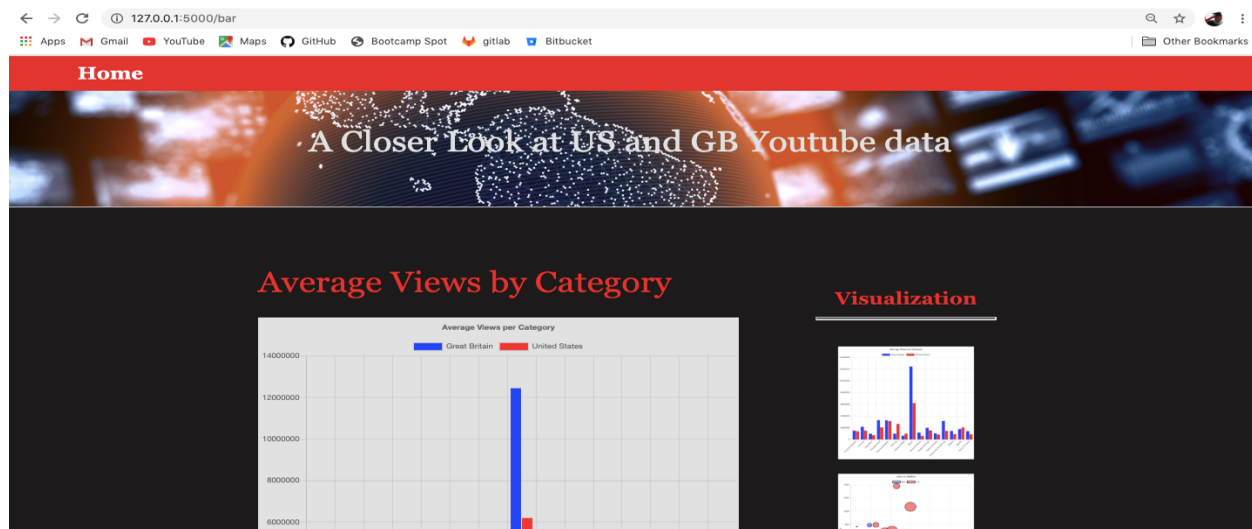
At the backend of this application are Python and Javascript. Python was used on the Extraction, Transfer and Loading process of the data to the database. It was also used to create the flask app that run this application. Javascript was utilized on rendering the visualization with the aid of the Chart JS library.

The dataset used was obtained from Kaggle.
It contains data for 10 countries with 41,000 entries each. The group decided to focus analysis on two countries - US and GB.

PostgreSQL was used on this project to store all the data obtained.
PostgreSQL is a powerful, open source object-relational database system that has earned it a strong reputation for reliability, feature robustness, and performance.
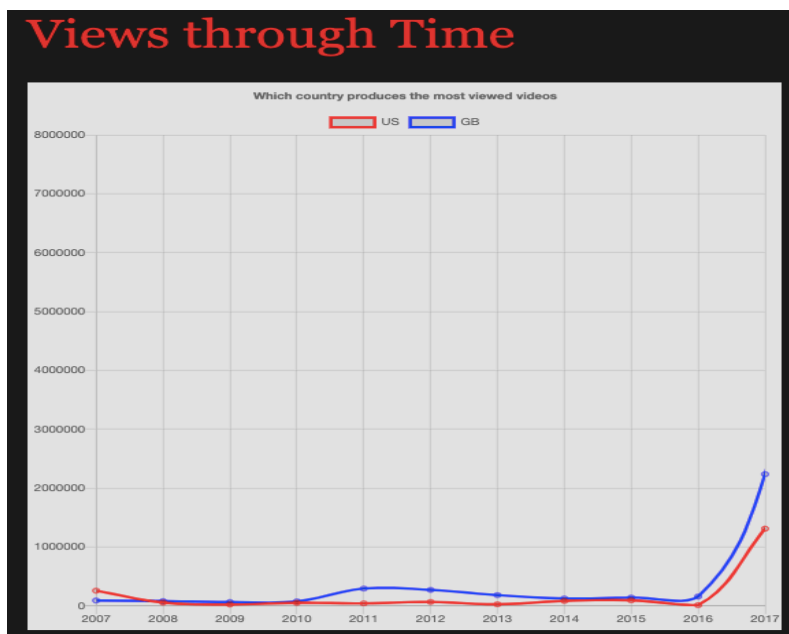
Description for View for Line Graph:
The line graph was created using chart.js library, A line chart is a way of plotting data points on a line. Often, it is used to show trend data, or the comparison of two data sets.

Demonstrates which country produces the most videos by total count.

The important parameters needed to render the output were:
the x values, shows the produced videos through a 10-year span, 2007 - 2017
the y values, provides the views by the Total number for all categories by country
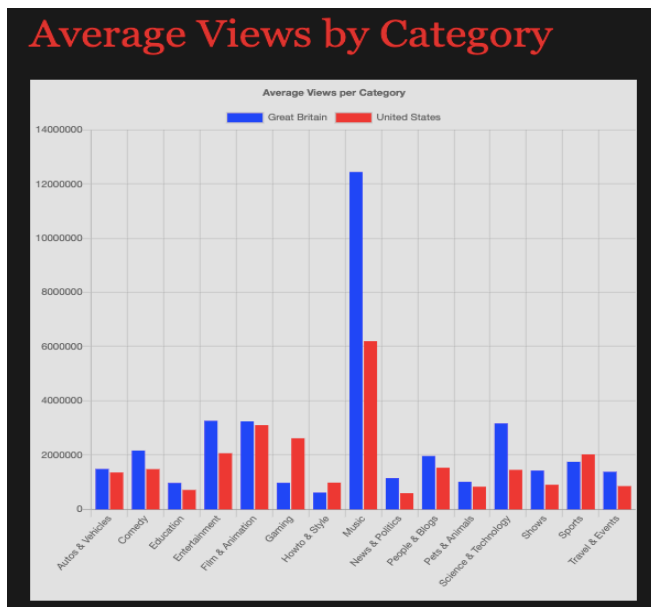
Description for View for Bar Graph:
The bar graph was created using chart.js library, bar chart provides a way of showing data values represented as vertical bars. It is sometimes used to show trend data, and the comparison of multiple data sets side by side.

Demonstrates Average Views by Category per country.

The important parameters needed to render the output were:
the x values, provides the views by the average totals
the y values, list the 15 categories by name

Description for View for Doughnut Graph:
The Doughnut graph was created using chart.js library, show proportions and percentages between categories, by dividing a circle into proportional segments. Each arc length represents a proportion of each category, while the full circle represents the total sum of all the data, equal to 100%.

Demonstrates Top 10 most liked categories per county, outer doughnut is GB, inner doughnut is US.
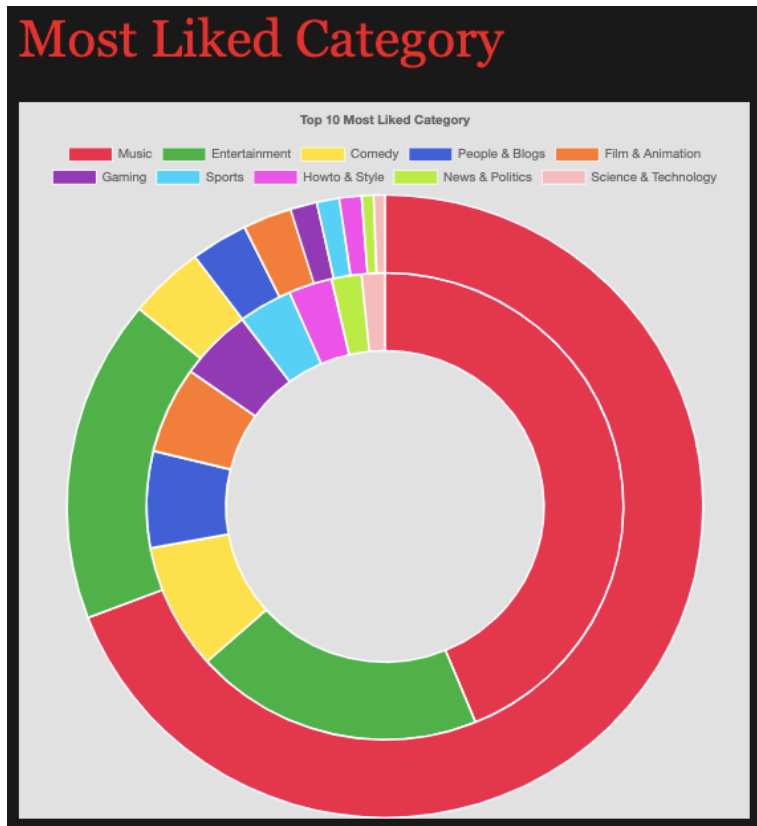
Description for View for Bubble Graph:
The bubble graph was created using chart.js library. A bubble chart is used to display three dimensions of data at the same time. The location of the bubble is determined by the first two dimensions and the corresponding horizontal and vertical axes. The third dimension is represented by the size of the individual bubbles.
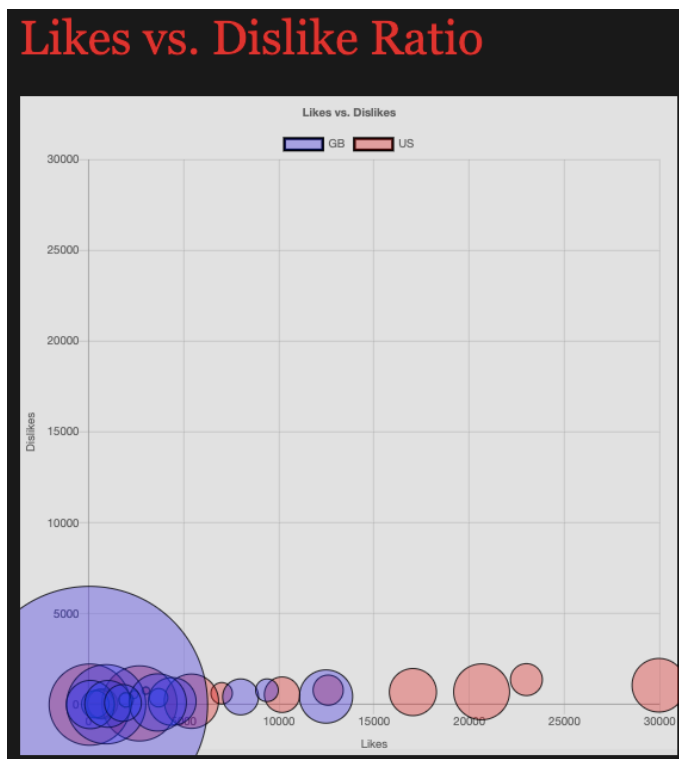
The important parameters needed to render the bubbles were:

the x values, which was set to the number of likes per category
the y values, which was set to the number of dislikes per category
the r values, which was set to the ratio of likes/dislikes per category
the labels, which was set to the categories.

5) Final conclusions and whether or not they met your hypotheses.
After completing our analysis based on how the data was sliced our hypotheses was not meet. The U.S first social media plat form is not You Tube. If you are looking to grow your business or create a trend, U.S first choice is Facebook. Furthermore, we were really surprised that GB number one social media platform is You Tube and that music is their most viewed and liked category. So, if you are looking for your next break through career in GB you might want to join the music industry. Now the next question is which arena and genre, we'll have to wait and see our next analysis.

6) Future Work
Include more countries.
Labels – Better labeling of Y and X axis
Doughnut Graph - update to % to show the difference between categories clearly
Y axis – Update the totals and average to be more consistent with data outputs