LEADING YOUTUBER'S ANALYSIS

Shivani Vellanki

Introduction:

The dataset chosen for the analysis is the leading two hundred youtuber's data. The data mainly focuses on the youtuber's followers(subscribers), likes, comments, the country to which they belong to, the category based on which they make the videos, the sub-category which explains about the category's group for example, if a youtuber's category is music and sub-category is music of Asia. The data also depicts about the core topic that was spoken more.

Problem Statement:

The problem statement for the chosen dataset analysis is

- Which geographical location has the greatest number of channels among the top 200 YouTube channels?
- For an aspiring or existing youtuber, which category of video must be chosen to acquire more views, subscribers, likes and average comments?
- To start a new channel in YouTube, what type of channel category must be produced to make the channel a sensation? These problem statements can be answered with the analysis of this dataset.

Number of YouTube channels from different countries:

There is an attribute named country in the dataset which represents the country to which the channel belongs to. The data is analyzed in such a way that the count of channels in a country is displayed. With this we can sort the data by concluding the **first problem statement.**


Sheet 1
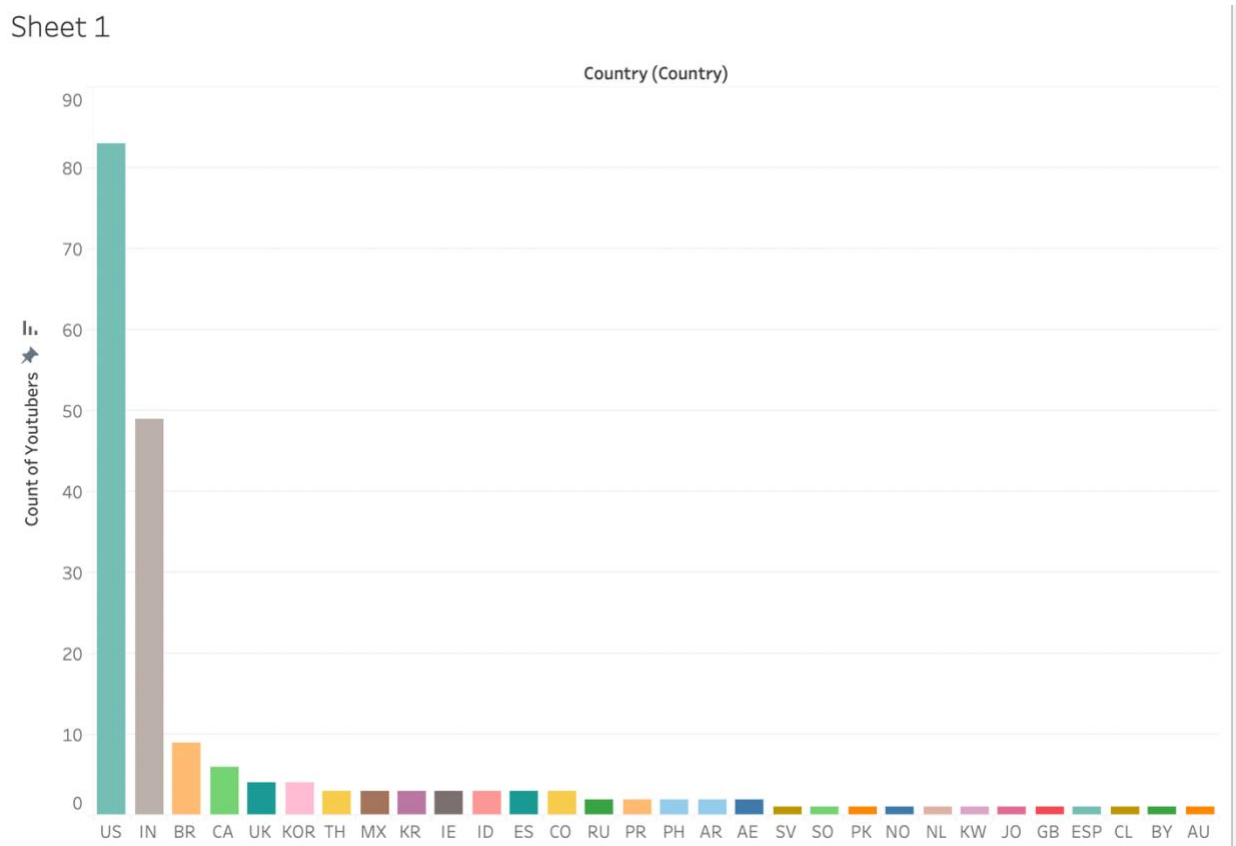
It is concluded that number of youtubers in United states are the highest. The probability of starting a channel and making it to the top from the USA is the maximum.

## Greater number of followers, interaction, and views for a category type:

The category type is the type of videos the youtubers make. There are fifteen different main video categories in the chosen dataset. For example, there is a main video category "Entertainment", youtubers make videos on entertainment like lifestyle, movies, and TV shows, these all being the sub-category. This analysis answers the **second problem statement** of what category videos must be made to gain more views and communication from the audience.

$\Rightarrow$ **Determining the number of Subscribers/followers according to the Video Categories:**

Main Video Category
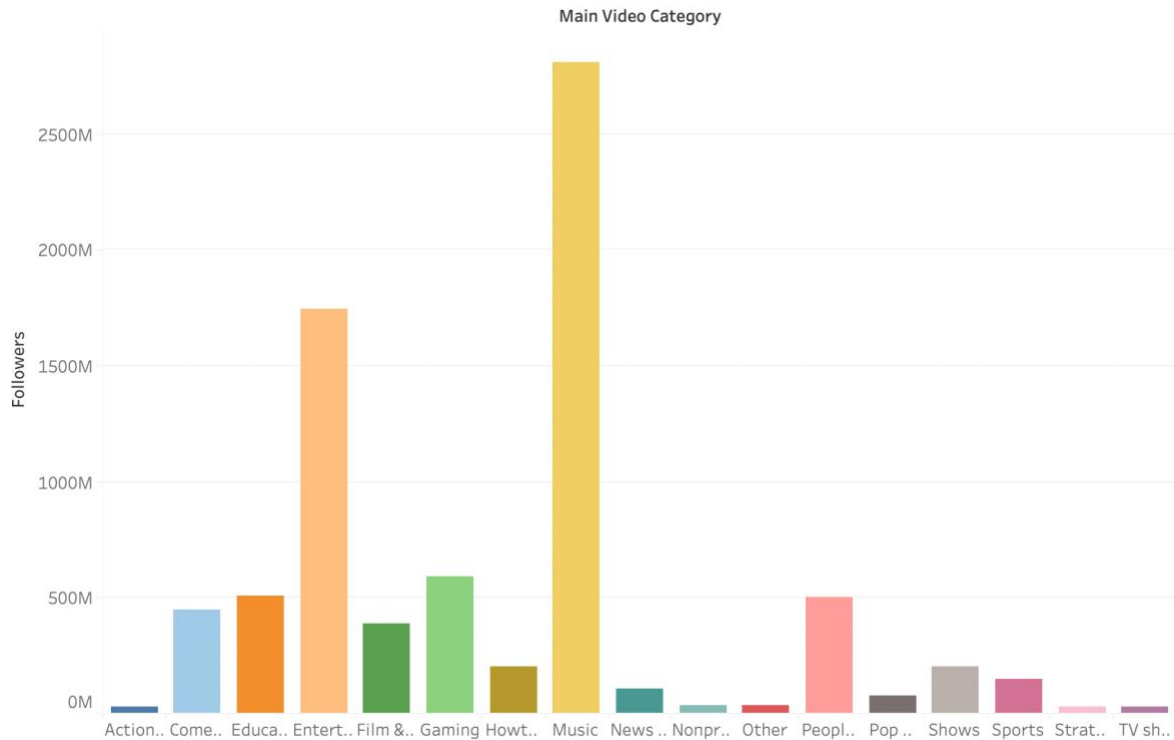
While analyzing the number of subscribers in the dataset, it is observed that maximum number of followers are for the music category, followed by the entertainment category, and then most of them follow gaming videos on YouTube.

⇒**Determining the number of likes according to the Video Categories:**

Main Video Category
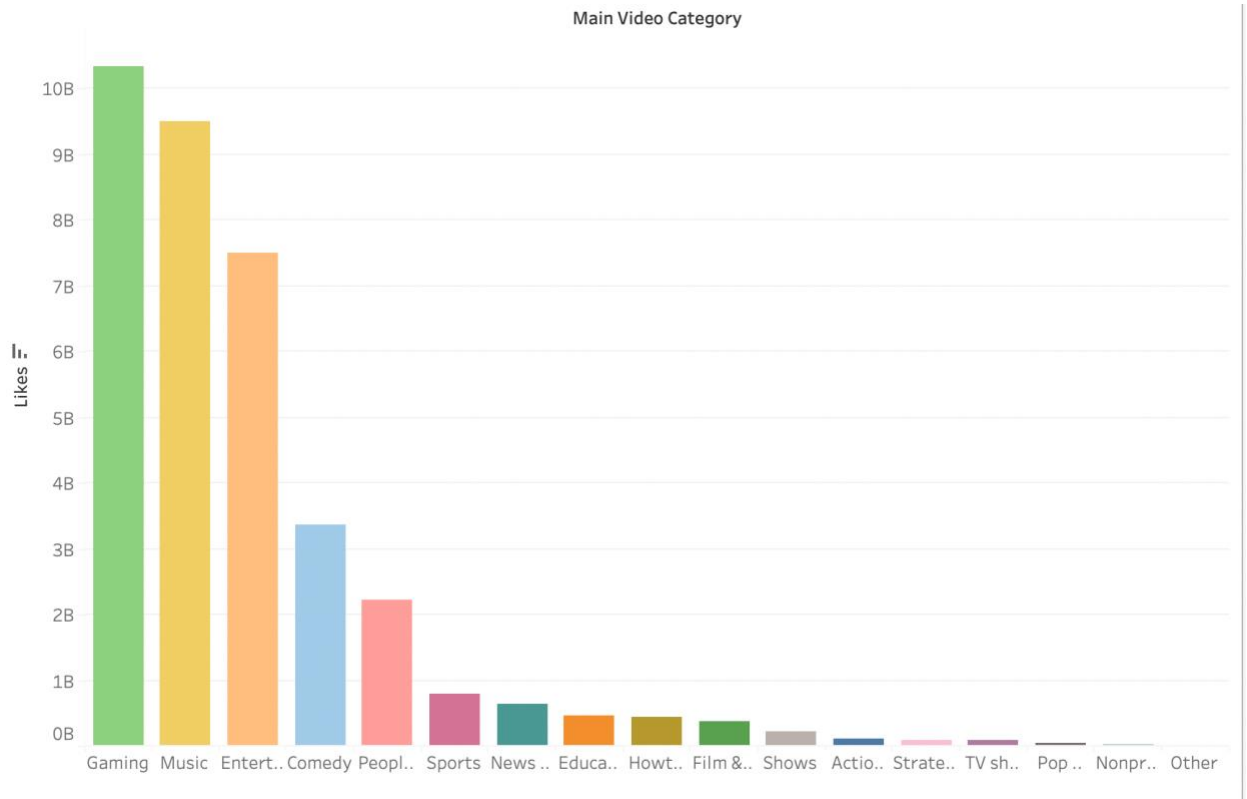
When observed the number of likes, most of the viewers like gaming videos, followed by music videos, and then the entertainment. It is noteworthy that, even though there are maximum number of followers for the music, number of likes are more for the gaming videos.

⇒**Determining the average number of comments according to the Video Categories:**
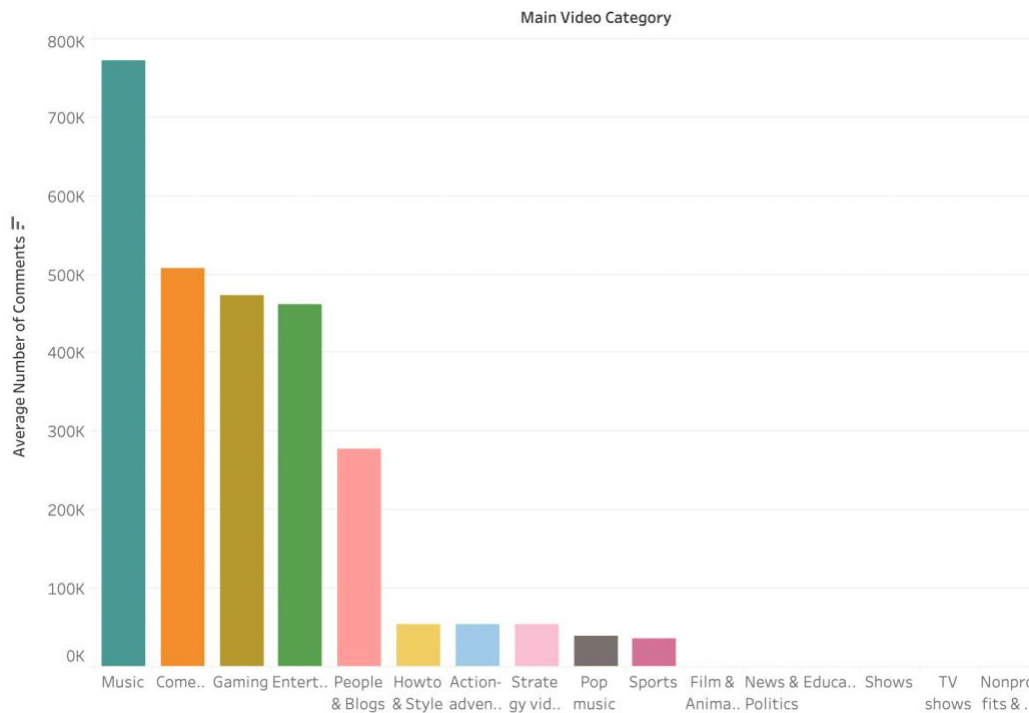
According to the dataset, the average number of comments are more for the Music category, further, intriguingly spectators' comment more on Comedy followed by Gaming and Entertainment.

⇒**Determining the number of views according to the categories:**

The number of views matters a lot for the video to go trending. The more trending the video gets, the more advertisement merchants collaborate which results in more money for each video.

Main Video Category
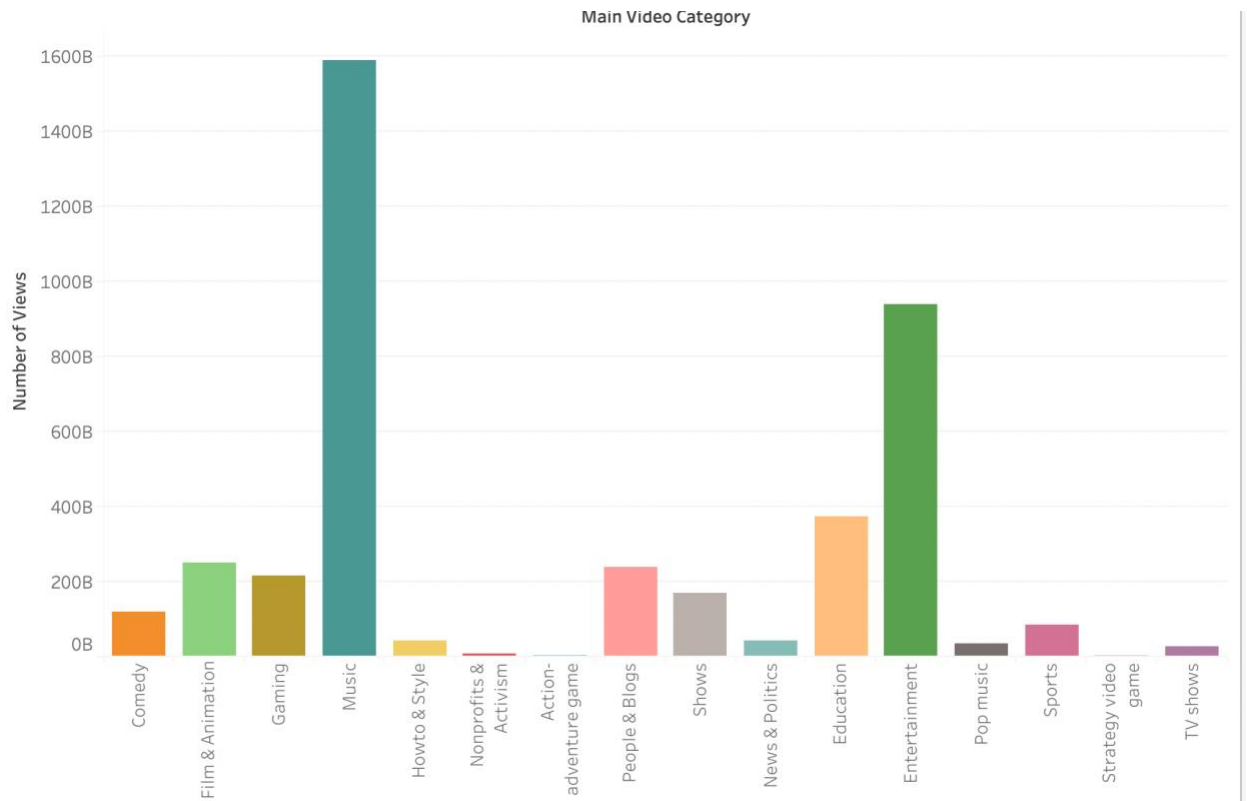
Audience preference on channel category:

There are two kinds of categories, channel, and the video category. The youtubers overall, make videos that are based on channel category, if it's the video category, the youtubers create videos subcategorized. This analysis answers **the third problem statement.**

Sheet 2

Category
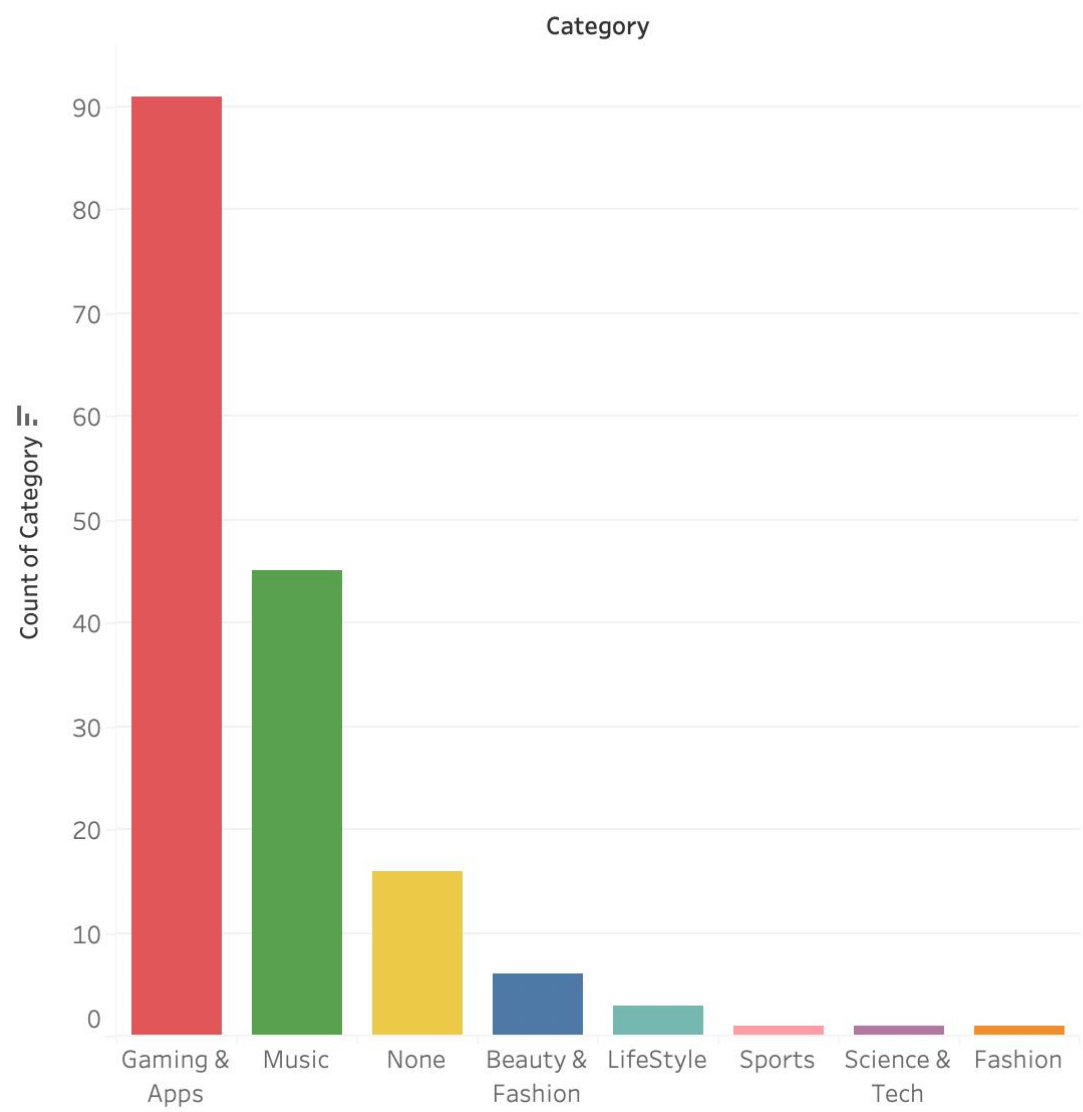
References:

Top 200 Youtubers Data (cleaned). (2022b, July 8).

Kaggle. https://www.kaggle.com/datasets/syedjaferk/top-

200-youtubers-cleaned