**TRENDING YOUTUBE STATISTICS ANALYSIS**

Founded in 2005, YouTube is arguable the biggest online video platform worldwide featuring a wide variety of user-generated and corporate media content that include music videos, TV clips, as well as other video clips such as video blogs, short original videos, Let's Play gaming videos, instructional videos on everything from language learning to stain removal, as well as videos with educational content.

**BACKGROUND TO THE PROBLEM**

**This** problem was issued to us as a group to study and analyze how youtube as a platform can help individuals, companies and other organization’s to understand the users view about a certain product in market or his/her suggestion(s) concerning a particular topic of discussion via a certain youtube channel. We are required to use an analytics tool called R and Rstudio to do our analysis and make conclusions based on what we are going to visualize.

Through this analysis, the company stakeholders can decide to either withdraw the product from market or continue with the product.

The datasets that we are going to use were extracted from some channels on youtube and we are required to analyze them based on a certain criterion. These datasets were collected using the youtube API.

Data is included for the US, GB, DE, CA, and FR regions (USA, Great Britain, Germany, Canada, and France, respectively), with up to 200 listed trending videos per day.

Through this problem, we shall be able to discuss the following in particular

1. Sentiment analysis in a variety of forms
2. Categorizing YouTube videos based on their comments and statistics.
3. Training ML algorithms like RNNs to generate their own YouTube comments.
4. Analyzing what factors affect how popular a YouTube video will be.
5. Statistical analysis over time

**THE PROBLEM THE PROJECT IS ADDRESSING**

To be able to use the analysis tool called R to understand what youtube does to determine the year’s top-trending videos, how youtube can help stakeholders to make decisions based on the analysis for example measuring user’s interactions (number of views, shares, comments and likes) and also to be able to use youtube to know the top performers on the YouTube trending list.

**THE MAIN GOAL OF THIS PROJECT**

To understand the basics of data analysis in R based on the data extracted from different youtube channels and how youtube does its data analysis and make decisions on what is the most trending video on youtube.