IBM HACK-CHALLENGE 2020

REPORT ON

COVID-19 tweets based Visualization Dashboard

Project ID: SPS_PRO_451

Team Members: Mansi Shinde

Sakshi Pophale

Simran Moondra

College Name: Cummins College of Engineering for Women

INDEX

1. INTRODUCTION

1.1 Overview

The Project is to develop a Visualization Dashboard - which shows the results of Sentiment Analysis on the Covid-19 tweets sent by Indians on the further Extension of the Lockdown. We have developed a Dashboard which displays the past 1 hour's Live tweets by Indians and presents the result of Sentiment Analysis in the form of Gauge chart and line graph too. The Gauge Chart shows the sentiment of live tweet and the line graph displays past 1 hour's live tweet sentiments. A tabular display also showcases the Top 5 tweets with Emojis(emoticons) alongside , to convey the sentiment behind the tweet.

Moreover, there are 2 separate tabs, where the former shows the sentiment analysis of 6 most affected States in the form of pie chart and line chart, and the latter shows the sentiment analysis of 3 most affected cities in the form of pie chart and line chart. Moreover, the important Updates/tweets on twitter given by World Health Organization and Ministry of Health Department can also be seen on the dashboard.

1.2 Purpose

Twitter is one of the social networking services where people can easily express their views related to any particular topic which are known as tweets. In this severe COVID-19 crisis, twitter will act as a best platform to understand what people think about the ongoing pandemic, government's decision on the lockdown extension.

Therefore, the purpose of this project is to perform the Sentiment Analysis on the tweets sent by the people during the further Extension of the Nation-wide and State-wide Lockdowns. The aim is also to build an Analytics Model and a Visualization Dashboard to understand behavior and sentiments of the people as the Lockdown is getting more and more extended.This would help the Government of India in taking better decisions by understanding what the citizens's reactions are towards the extension of lockdown.

2. LITERATURE SURVEY

2.1 Existing problem

It is very crucial for the Government to understand the Sentiments and reactions of the country before taking any further decisions and making any new announcements. They should also know the emotions of the Nation and the problems the people have been facing in the existing Lockdown . So, that in the next announcement of Lockdown , the authorities knowing and deducing the citizen's emotional state can amend our policies beneficial to them. Also, e-commerce stakeholders can adjust according to their state and regulate products demand and supply.

2.2 Proposed solution

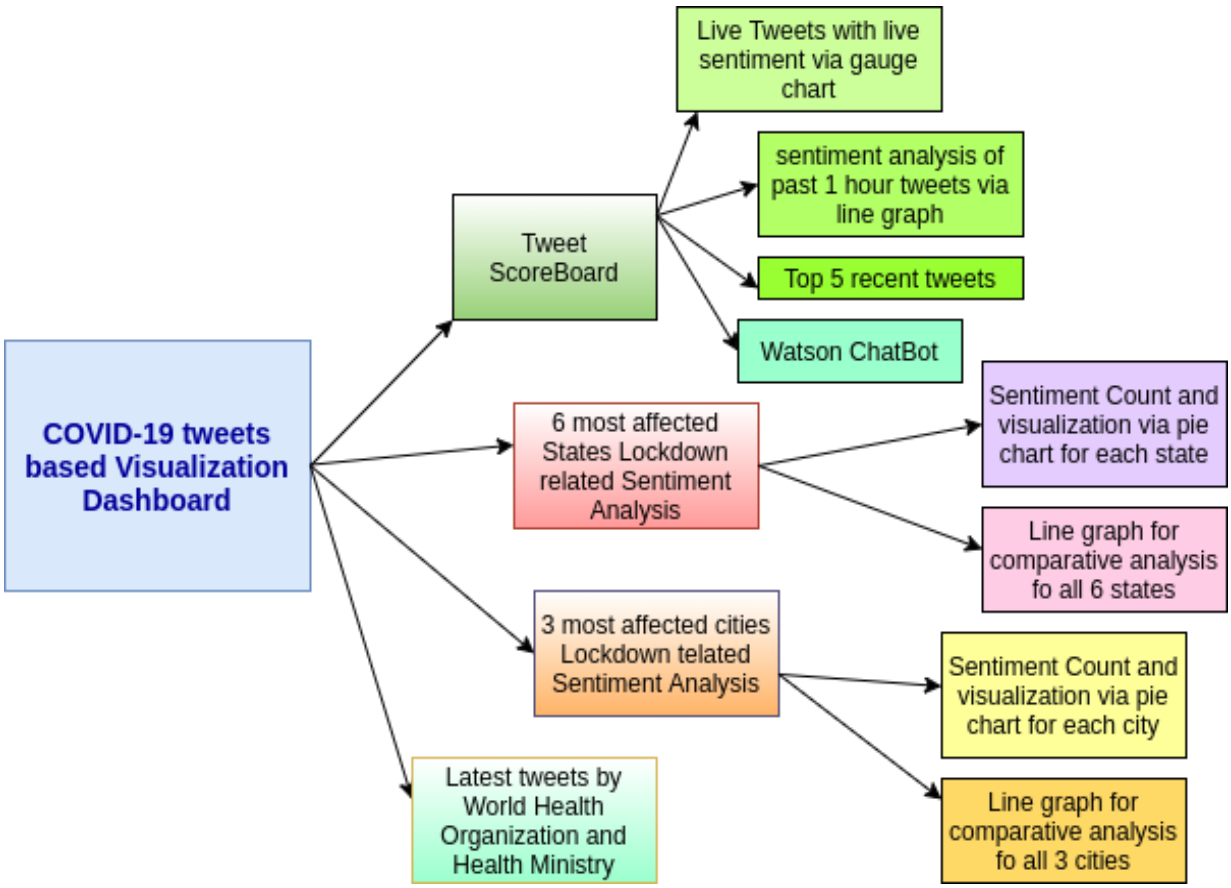
Twitter is one of the social networking and microblogging service where people can easily express their views related to any particular topic which are known as tweets. In this severe COVID-19 crisis, twitter will act as a best platform to understand what people think about the ongoing pandemic, government's decision on the lockdown extension.

We put forward the solution how these tweets of people related to the current COVID crisis can be used to understand their sentiments in a better way and can be visualized in the form of a dashboard. It will also help to predict people's reaction to the government's decision on extension of lockdown. The

graph generated by the dashboard will help the government to take better decisions during this pandemic.

3. THEORETICAL ANALYSIS

3.1 BLOCK DIAGRAM



3.2 SOFTWARE DESIGNING

There were various software/application design decisions made before arriving at the final solution.

Earlier it was proposed that Python's tweepy library would be used to extract the tweets and then use python's sentiment analysis libraries like Vader or Textblob to perform sentiment analysis on the extracted tweets. However, later after studying and doing some hands-on on Node Red on IBM cloud, building Dashboard using Node Red was found quite easier and efficient as all the nodes related to tweets extraction, sentiment analysis, charts etc that were needed to design a Dashboard were already present or were just needed to install their respective packages. Moreover, Node Red is very user friendly, anyone from a non coding background can start using it from scratch and develop applications.

The nodes used in this project are - Twitter nodes, Sentiment Analysis nodes,Cloudant nodes, ui-table node, function nodes etc

Three Cloudant databases are used for this project. All three are used to store the Live Tweets, so as to perform Sentiment Analysis on them and produce the counts of Positive, Negative and Neutral Tweets and then further represent them in Pie Chart and Line chart. Storing the Live tweets also helps in showcasing the Top 5 tweets in the table on the Dashboard.

The First Cloudant Database is used to store live tweets coming from all over India. These Live tweets are displayed along with its sentiments using a gauge

chart. There is a line chart which displays past 1 hour's live tweets. Here, the cloudant is used to get the sentiment counts of past 1-2 days tweets along with live tweets, i.e the live tweets keep on getting stored on cloudant and the results on pie chart will be updated after every 10 minutes. This also helps to showcase top 5 recent tweets stored on cloudant. There is also a flow, which will delete the old tweets, but it has to be triggered by the user as and when needed as a sign of maintenance of the dashboard.

The Second Cloudant Database is used to store the live tweets of 6 most affected States - Maharashtra, Delhi, Tamil Nadu, Gujarat, Uttar Pradesh and Karnataka, so as to perform Sentiment Analysis and plot the Pie-chart of the Counts of Positive, Negative and Neutral tweets of the respective states and line chart to display the comparative analysis of counts of all 6 states. There is also a flow, which will delete the old tweets, but it has to be triggered by the user as and when needed as a sign of maintenance of the dashboard.

The Third Cloudant Database is used to store the live tweets of 3 most affected cities - Mumbai, Bangalore and Chennai, so as to perform Sentiment Analysis and plot the Pie-chart of the Counts of Positive, Negative and Neutral tweets of the respective states and line chart to display the comparative analysis of counts of all 3 cities. There is also a flow, which will delete the old tweets, but it has to be triggered by the user as and when needed as a sign of maintenance of the dashboard.

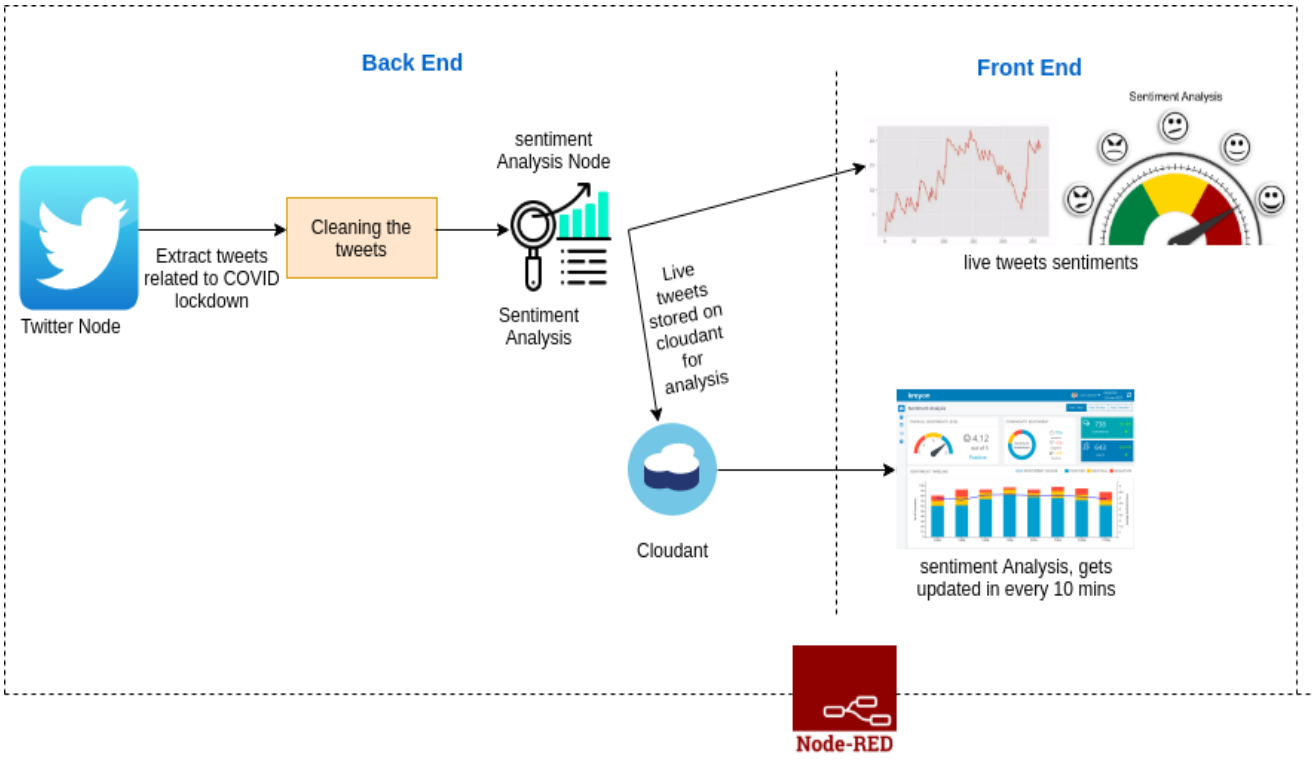
4. EXPERIMENTAL INVESTIGATIONS

In the process of developing the project, we underwent many investigation processes, in order to build this Dashboard successfully. Initially, investigations were made, how python's tweepy could be used to extract tweets and use python's sentiment libraries and plots to build a dashboard. But as it was decided by our team that we are going to use all the services provided by IBM cloud to build our project, building user interactive dashboard would be complex and would require lots of investigations and integrations. Node Red was an easy tool to build interactive dashboards, as it had node which supported most of the services for example,

1. Twitter Node - to extract live tweets related to a particular hashtag, or user.
2. Sentiment Analysis Node - to predict the sentiment of any tweet or text. It contains the resulting AFINN-111 sentiment.
3. Cloudant Node - to insert, update or delete the tweets
4. Template Node - To build the UI, insert Images, etc
5. Chart Node - to build analytical charts/graphs for visualization of analysis
6. Watson ChatBot Node - to include the chatbot on dashboard to resolve user queries.

The project was completely based on experimenting with the Node Red Tool, where we code the application in the form of building the flow of that application on the Node Red Canvas.

5. FLOWCHART


Back End
Front End
Twitter Node
Extract tweets related to COVID lockdown
Cleaning the tweets
sentiment Analysis Node
Sentiment Analysis
Live tweets stored on cloudant for analysis
Cloudant
Sentiment Analysis
live tweets sentiments
sentiment Analysis, gets updated in every 10 mins
Node-RED

6. RESULT

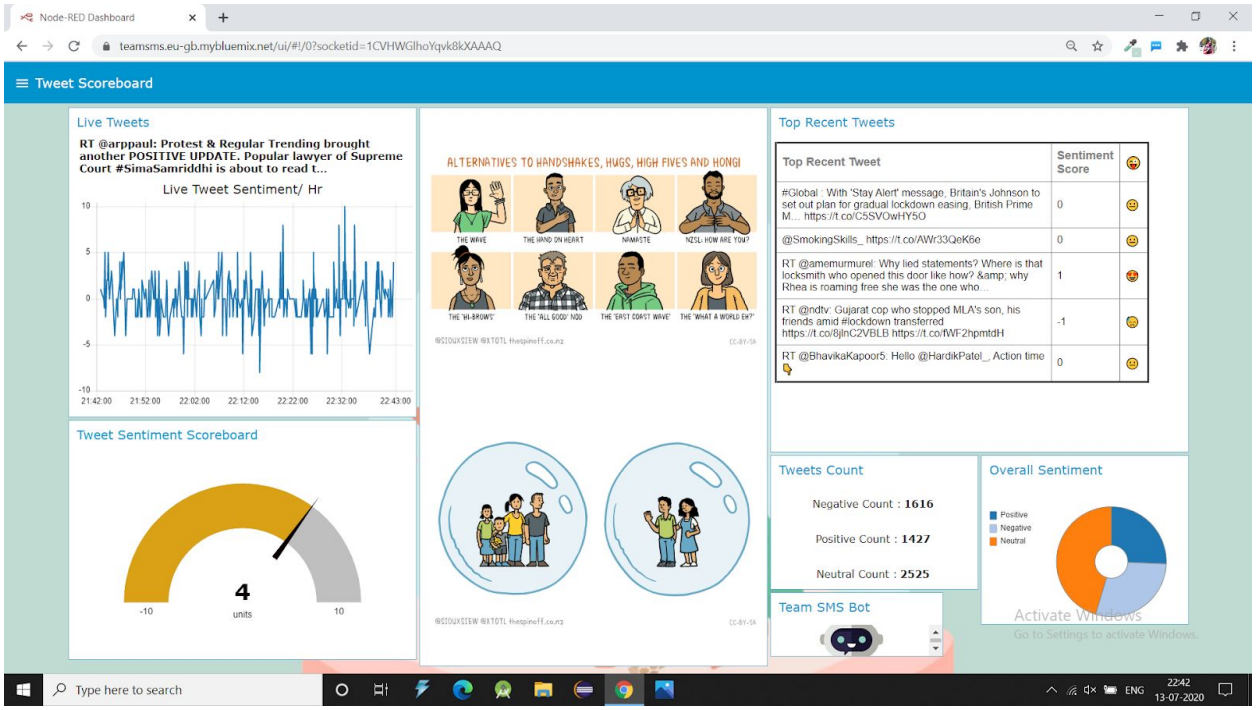


Fig : The Tweet Scoreboard tab.

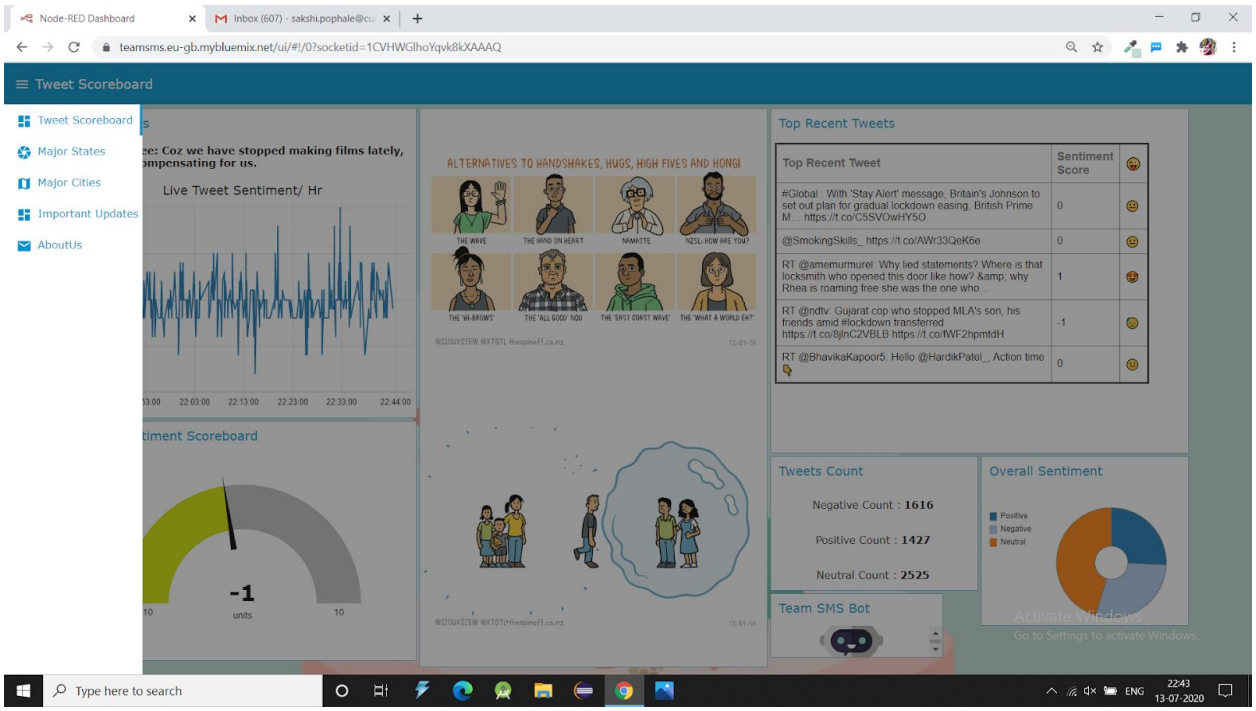


Fig : The Menu bar for Navigation

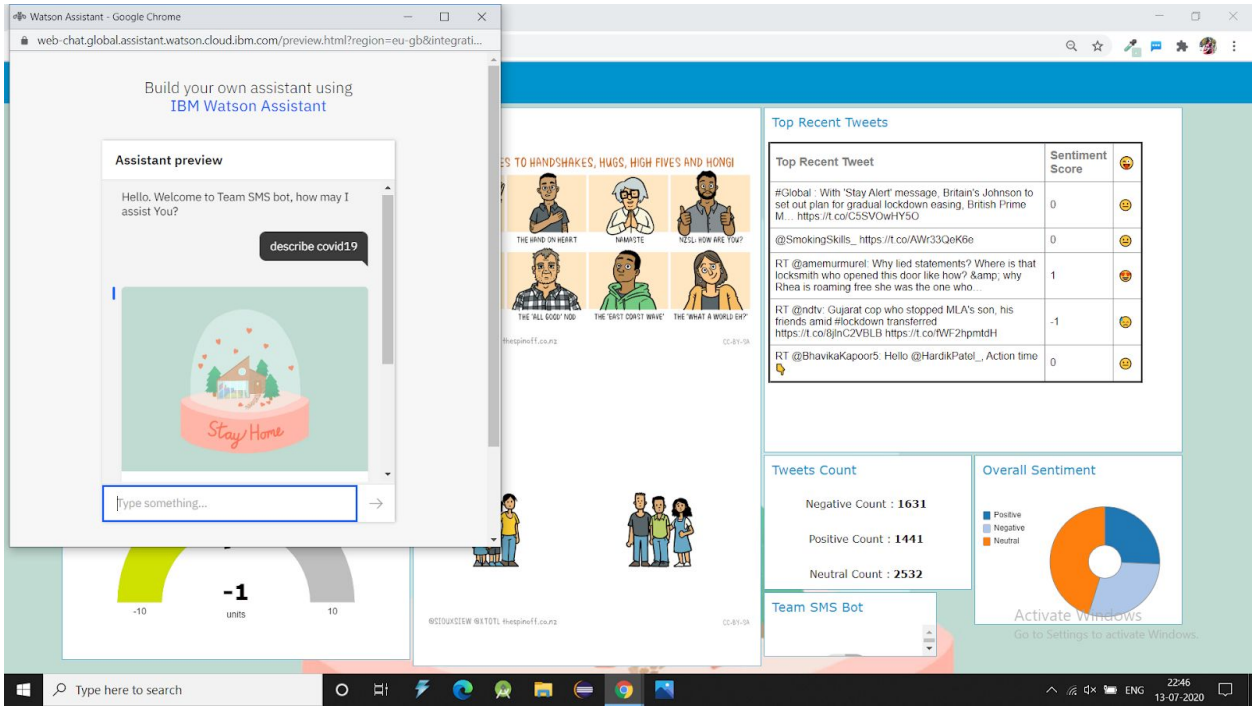


Fig : The Chat-Bot Window

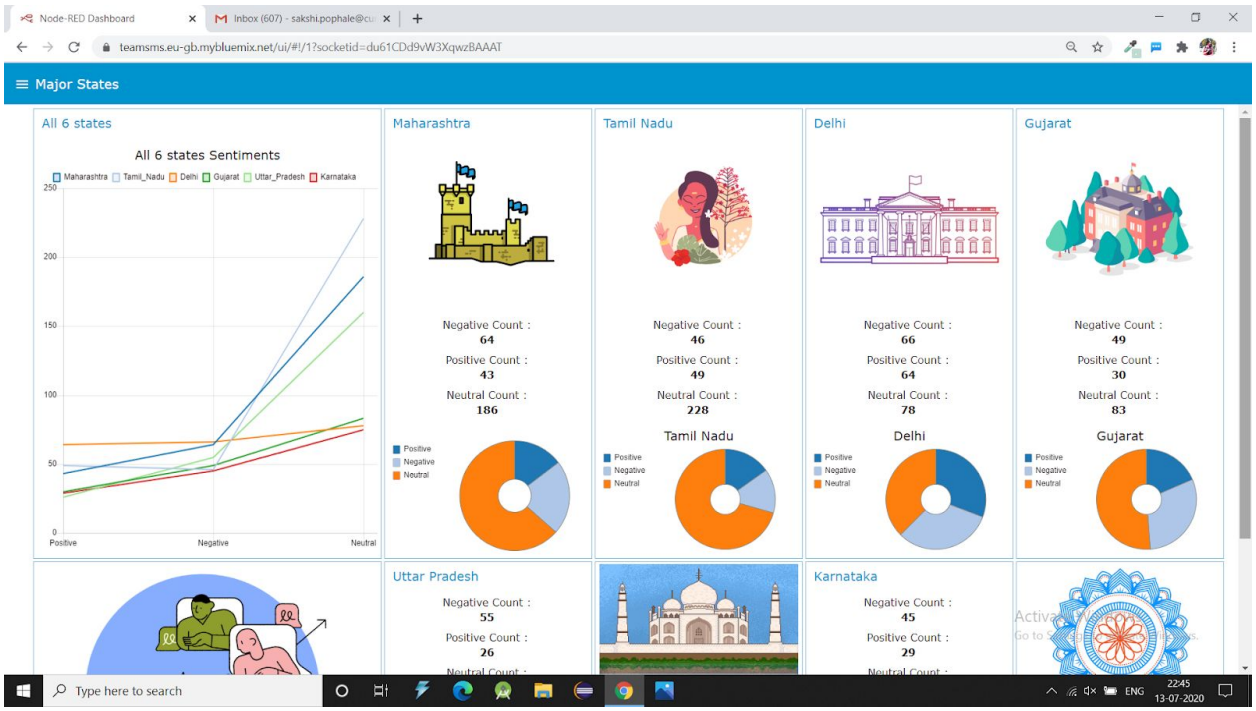


Fig : The Sentiment Analysis Counts for the most affected States.

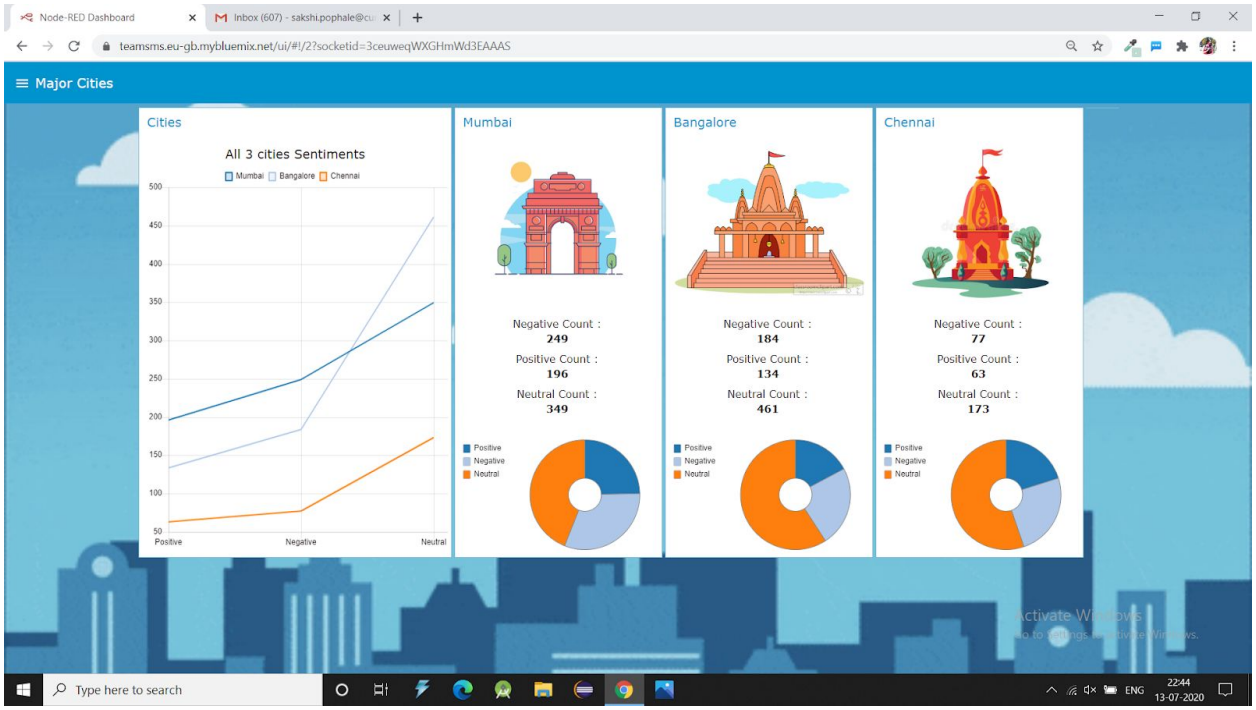


Fig : The Sentiment Analysis Counts for the most affected Cities.

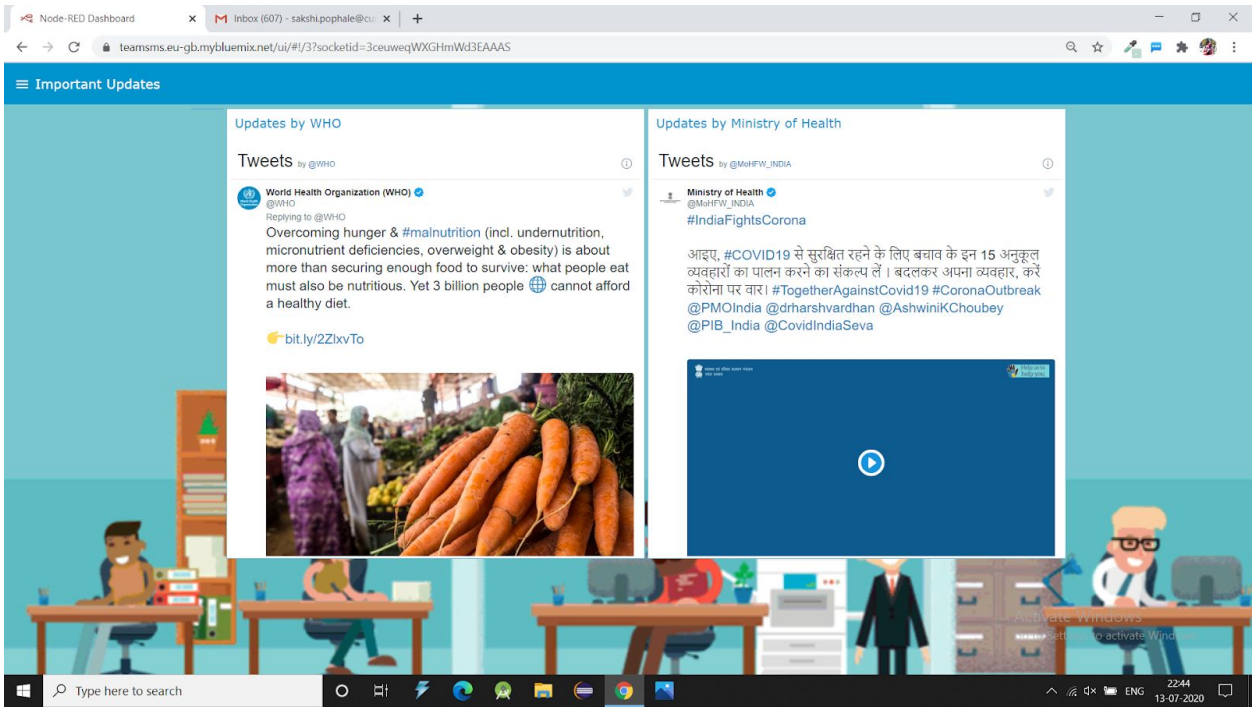


Fig : The Important Updates tab (tweets by WHO and Ministry of Health)

7. ADVANTAGES & DISADVANTAGES

7.1 ADVANTAGES

The advantages are as follows :-

i. This Visualization Board is very advantageous to the Government Authorities as it will help them understand the Nation's reactions to their Announcements.

ii. It will further help the Government to amend their decisions for the next lockdown's guidelines or any new announcement.

iii. The Dashboard also helps the Government and Researchers and other concerned Authorities understand the emotions of the citizens and how they can contribute in helping the people and motivating them in such trying times.

iv. Also, e-commerce stakeholders can adjust according to their state and regulate products demand and supply.

7.2 DISADVANTAGES

The disadvantages are as follows:

i. While social media cannot reflect the sentiment of the total population of any country, it can surely be counted as a sample population to determine the vibe of a nation.

ii. Sentiment Analysis algorithms have a problem in understanding Sarcasm and Irony in sentences.

iii. The twitter API will not deliver 100% of all the tweets.

8. APPLICATIONS

The Applications are as follows :-

- Gives a big help to the Government in understanding their citizens's thinking and emotions and problems, so as to help them in making a better decision for further announcements.
- E-commerce stakeholders can adjust their products requirements according to the demands and regulate products demand and supply.
- It even helps Researchers , especially Psychologists to understand and study people's reactions and behavioral changes in such testing times.

9. CONCLUSION

The main aim of this project is fulfilled by the creation of the Visualization Dashboard and showcase the Sentiment Analysis Scores (in the form of various Charts). With the world going through a turbulent situation because of the COVID 19 pandemic, it is crucial to see that the lockdown and other restrictions do not impact the psychological well-being of the population.

The live tweets present on the Dashboard are very helpful to voice the thinking of the people directly to the government decisions. Such Analysis results should be very helpful to the Government Authorities, as they make further decisions in such trying times.

10. FUTURE SCOPE

For future additions to the Dashboard , presentation of the latest news updates on Covid-19 from various news sources. Give necessary home remedies for boosting immunity. To present light upon the facts and myths on COVID19.

11.BIBILOGRAPHY

Team Members:

1. Mansi Shinde
2. Sakshi Pophale
3. Simran Moondra

College Name: Cummins College of Engineering for Women

Project Name: COVID-19 tweets based Visualization Dashboard.

References:

1. https://flows.nodered.org/node/node-red-node-twitter

2.https://nodered.org/docs/user-guide/writing-functions#getset-multiple-values

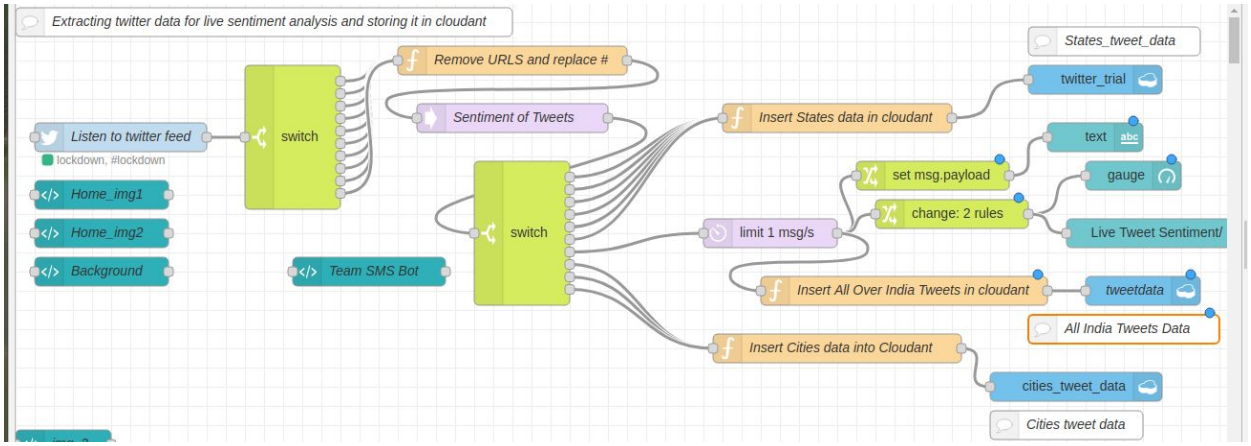
3.https://github.com/johnwalicki/Node-RED-Twitter-Workshop

4.https://developer.ibm.com/patterns/analyze-twitter-handles-and-hashtags-for-sentiment-and-content/

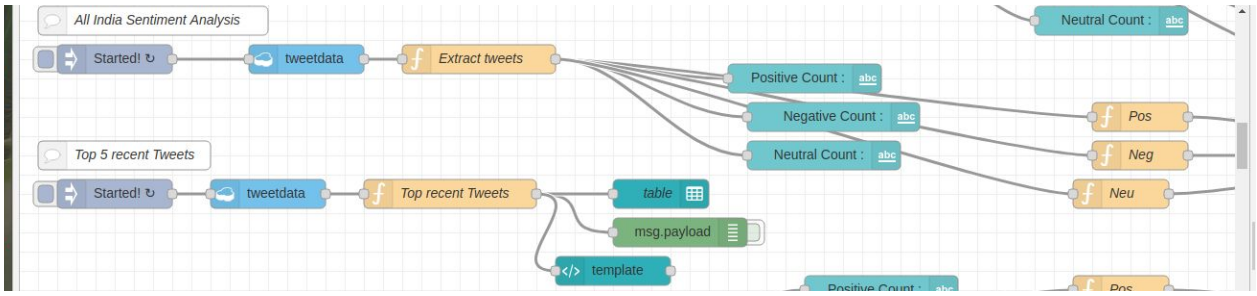
APPENDIX

A. Source code

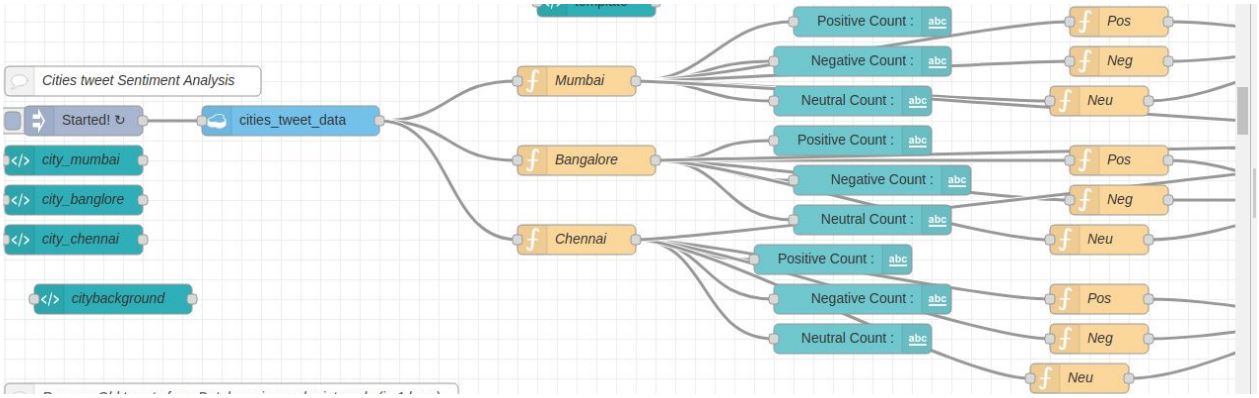
1. Tweets extraction, cleaning and getting sentiments



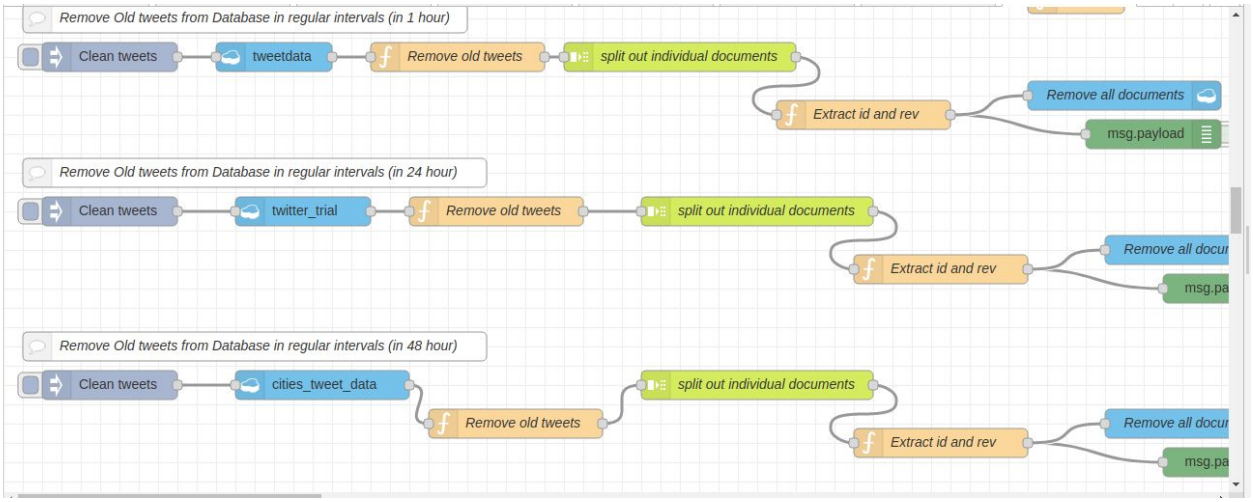
2. Sentiment Analysis and top 5 recent tweets



3. Cities sentiment analysis



4. Deleting old tweets from cloudant



Node Red Flow Link:
https://teamsms.eu-gb.mybluemix.net/red/#flow/c5065f3a.48774

Node Red DashBoard Link:
https://teamsms.eu-gb.mybluemix.net/ui/#!/0?socketid=8MzV24LaifwIXTOoAAAD

Youtube Video Link:
https://youtu.be/wlh0sCmyJKY