**Abstract**

COVID-19 (Corona Virus Disease 2019) has significantly resulted in a large number of psychological consequences. The aim of this study is to explore the impacts of COVID-19 on people’s mental health, to assist policy makers to develop actionable policies, and help clinical practitioners (e.g., social workers, psychiatrists, and psychologists) provide timely services to affected populations.

For this project we extracted and analyzed tweets of 10,000 active users of twitter from the period just before the announcement of first lockdown (w.e.f. 18th March to 20th March) and 10,000 tweets from just after the first lockdown (w.e.f. 25th March -27th March).

We calculated word frequency, scores of emotional indicators (e.g., anxiety, depression, indignation, and happiness) from the collected data. The sentiment analysis and the paired sample t-test were performed to examine the differences in the same group before and after the declaration of all India lockdown on 18 March, 2020.

**Extracting the Tweets**

Our team extracted 10000 tweets about coronavirus and related words for each timeframe, before and after lockdown, by using the Tweepy API in a python script.

The dataset for each of the periods, before and after, was extracted independent of each other and then stored in a CSV format for suitable analysis.

Data Cleaning and Tokenization

First, the text column is pre-processed with making all characters lowercase, removing all punctuation marks, white spaces, and common words (stop words), usernames associated with tweets and storing the cleaned tweets in a column named tidy\_tweets in the Data Frame.

After pre-processing, we used the tokenization method to grab the word combinations in the two datasets.

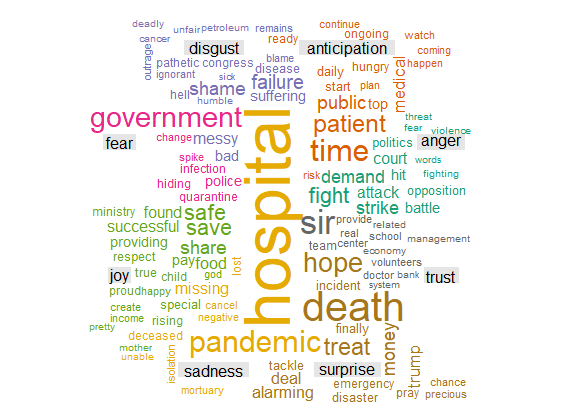
**Word frequency analysis**

One of the easiest ways to explore data is with frequency analysis. Although not difficult, in sentiment analysis this simple method can be surprisingly illuminating.

Top five words with highest frequency (before lockdown):

1. India (5572 times)
2. Corona (1050 times)
3. Narendramodi (1003 times)
4. Cases (952 times)
5. People (929 times)

**Sentiment Word cloud**



Top five words with highest frequency (after lockdown):

1. India( 8304 times)
2. Lockdown ( 1662 times)
3. Cases (1389 times)
4. Narendramodi (723 times)
5. Virus (718 times)

**Sentiment Word cloud for after lockdown**



## Sentiment word frequency

Next, we can also categorize words using polarity scores, putting the words into the ‘positive’ or ‘negative’ basket. First, we assign each word in the extracted tweets a polarity score. Then, we filter the dataset to get only words with a polarity score of 80 or more.

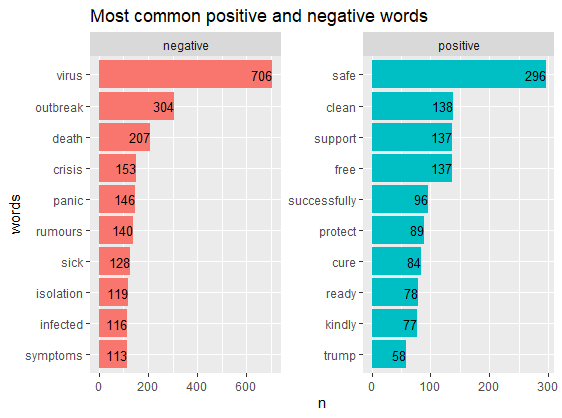
## 

## 

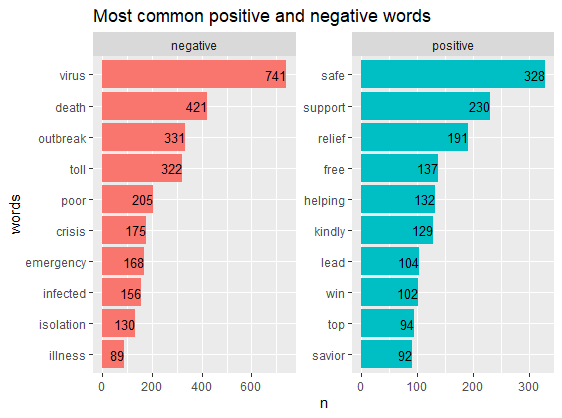
## Most positive and negative words used in the extracted tweets

To get a broader picture of how positive and negative words are used, we assigned the words with a sentiment using the ‘bing’ lexicon and did a simple count to generate the top 10 most common positive and negative words used in the extracted tweets.

**Before the lockdown**



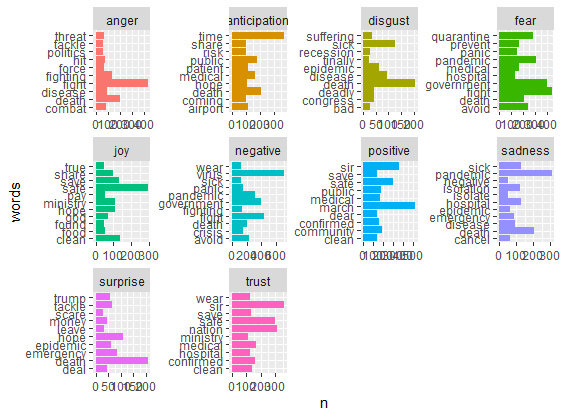
**After Lockdown**



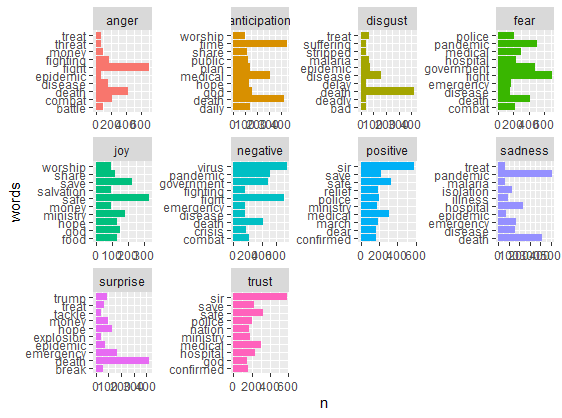
## Categorizing the sentiments into ten types of emotions

Other than categorizing the words into two categories (positive and negative) only, we can also label the words into multiple emotional states. Here we have grouped the words into ten groups and pick the top 10 most common words in each of the group.

**Before Lockdown**



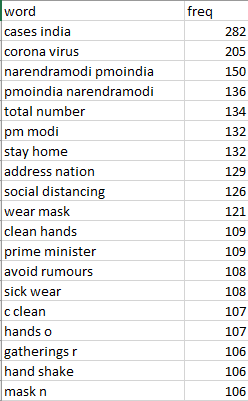
After Lockdown



**Bigrams and Trigrams**

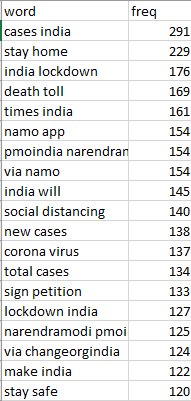
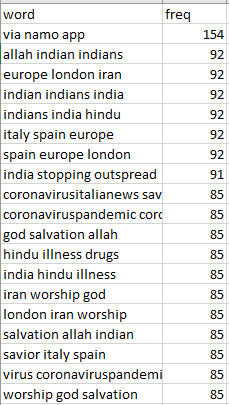
Word frequency of bi-grams (2-word) and tri-grams (3-word) can give us better insights from the dataset we are analyzing.

Insights from the same have been enlisted below:

According to the bigram and trigram for before lockdown period people are talking mostly about:

1. Number of cases in India.
2. Wearing masks and taking the required precautions against Covid-19.
3. Coronavirus.
4. Prime Minister Narendra Modi and the policies of his government towards the pandemic.

According to the bigram and trigram for after the lockdown people are mostly talking about:

1. Number of cases in India.
2. Aarogya Setu App.
3. Staying at home.
4. Lockdown.