Sentiment Analysis Of COVID-19 Tweets – Visualization Dashboard

Team Name : BitCoders

Team Members :

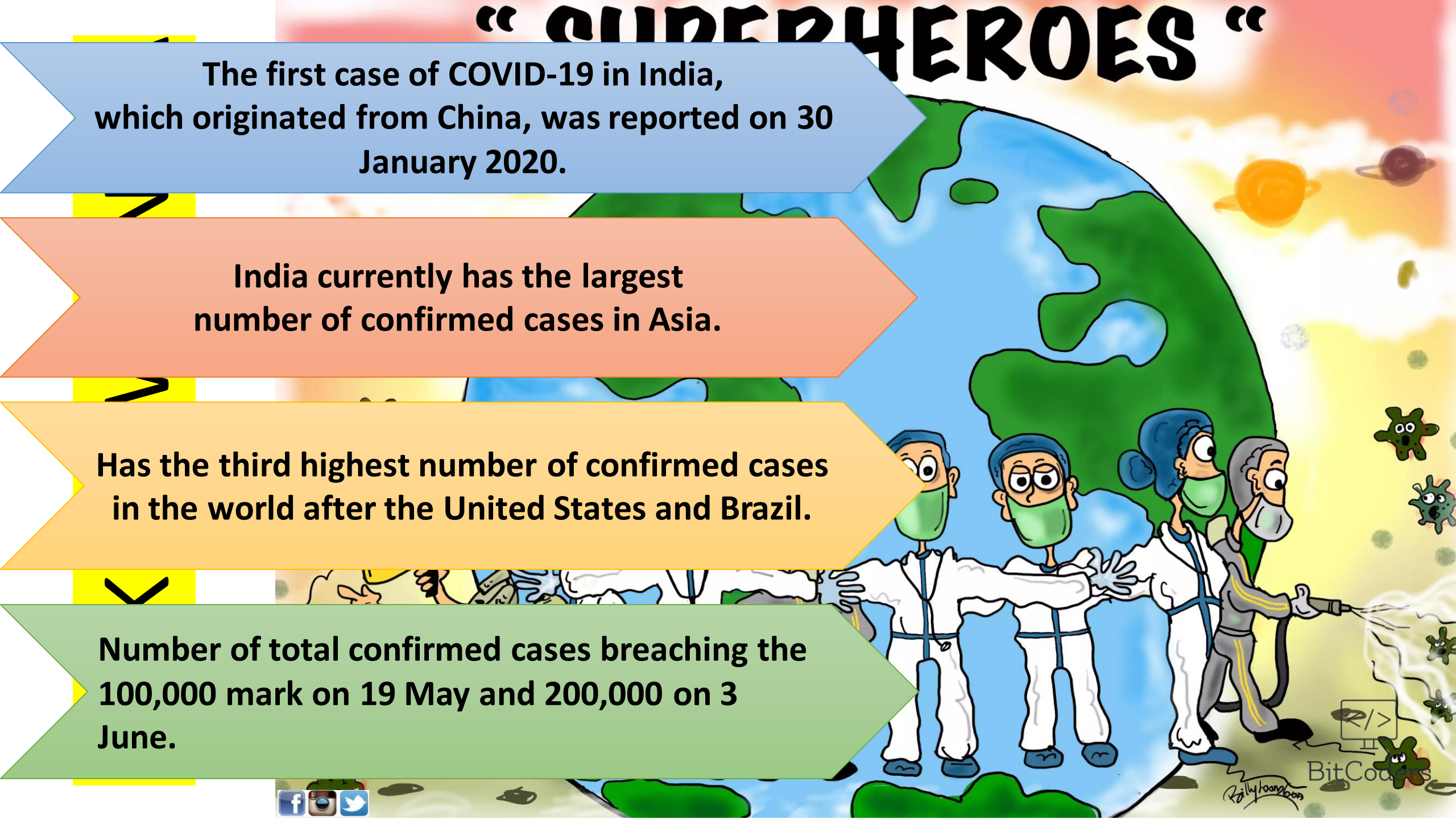
Himanshi Gupta (Team Lead)

Abhirpiya Sharma

Shruti Mishra

BitCoders

LOCKDOWN
STAY HOME


" SUPERHEROES "
The first case of COVID-19 in India, which originated from China, was reported on 30 January 2020.
India currently has the largest number of confirmed cases in Asia.
Has the third highest number of confirmed cases in the world after the United States and Brazil.
Number of total confirmed cases breaching the 100,000 mark on 19 May and 200,000 on 3 June.
BitCoders

7:54 AM 4G

Tweet

World Health Org
@WHO

Preliminary inves
conducted by the
authorities have
clear evidence of
to-human transm
of the novel #cor
(2019-nCoV) ider
#Wuhan, #China

CORONAVI
(2019-nC

Tweet your reply

Arvind Kejriwal
@ArvindKejriwal

PM has taken correct decision to extend lockd
many de

Sharad Pawar
@PawarSpeaks

I had a discussion with Hon Chief Minister of
Maharashtra Shri @OfficeofUT regarding cur
#Covid_19 situation in state & the challenge
the administration & preventive measures to
relief to various sections. I have conveyed my
suggestions on following topics.

10:32 AM · May 20, 2020 · Twitter for Android

224 Retweets **1.7K** Likes

Sharad Pawar @PawarSpeaks · 4h
Replying to @PawarSpeaks
The next academic year will be delayed due to the situation
and lockdown. As a result, the number of students & teache
The income of educational institutions & technological insti
get adversely affected.
@CMOMaharashtra
#Educacion

11 68 330

Sharad Pawar @PawarSpeaks · 4h
Some educational institutions are likely to collapse or close
financial losses. A study group or committee should be app
timely measures to ensure that students,teachers & educati
are not harmed & process of education is not disrupted.

11 61 273

Sharad Pawar @PawarSpeaks · 4h
The government has emphasized starting industries in the s
the conditions of lockdown. However, the guidelines for the
execution for the same are not comprehensive enough.

5 55 246

Press Trust of India
@PTI_News

Thousands of devotees gather in temples in various parts of West Bengal on occasion of Ram Navami defying lockdown imposed across country due to #coronavirusoutbreak, police ask them to return home immediately

4:50 PM · 02 Apr 20 · PTI_Tweets

1,181 Retweets **2,701** Likes

Failed to load Tweet. RETRY

Pinned Tweet

अनुज बाजपेई
@Real_Anuj

...
ायरस से भी भयंकर है,
ायरस !

0 करोड से भी अधिक संक्रमित !

'weet

Sagar
@sagarcasm

akele rehte hain unko family ki yaad aa rahi hai.
family ke saath rehte hain unko rona aa raha hai.

navirusLockdown

weet
Mar 24, 2020 · Twitter for Android

eets **4.2K** Likes

Sentiment Analysis

Sentiment Analysis

Twitter

PROBLEM STATEMENT

Description : The sentiment analysis of Indians after the extension of lockdown announcements to be analyzed with the relevant #tags on twitter and build a predictive analytics model to understand the behavior of people if the lockdown is further extended.

Also develop a dashboard with visualization of people's reaction to the government announcements on lockdown extension.

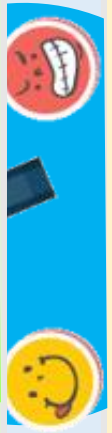
Expected Solution : Develop a twitter sentiment analysis model to understand the following

1. Get to know people's sentiment towards the epidemic
2. Understand the sentiments of people in the government. decision to extend the lockdown.

BitCoders

Twitter Sentiment Analysis

Sentiment Analysis



330 million monthly active users

145 million daily active users

Pull-able

Data

Machine Learning Technique

Detects Polarity

Interpretation and Classification of emotions within text data

Positive, Negative and Neutral

Twitter Sentiment Analysis

Twitter Sentiment Analysis

customers are able to express their thoughts and feelings more openly than ever before

Real-Time Analysis

Scalability

Consistent Criteria

TIMELINE

June 6 – June 15

Project Kick-off - Understanding the problem statement.

June 15 – June 20

Data Collection and Cleaning

June 21 – June 25

Data Analysis and finding new approaches

June 26 - June 30

Sentiment Analysis And Visualization

July 1 – July 5

Dashboard creation

July 6-July 15

Project Reports , video, and ppt creation

BitCoders

Technologies and Framework Used

USER

IBM Watson Dashboard

CLOUD

IBM WATSON

IBM

jupyter

EXTERNAL FOR DATA COLLECTION

kaggle

Solution Approach

1 Data Collection

Used kaggle to extract the Dataset containing tweets for March 29, 2020. It contains 22 columns 564141 rows.

'user_id', 'created_at', 'screen_name', 'text', 'source', 'reply_to_status_id', 'reply_to_user_id', 'reply_to_screen_name', 'is_quote', 'is_retweet', 'favourites_count', 'retweet_count', 'country_code', 'place_full_name', 'place_type', 'followers_count', 'friends_count', 'account_lang', 'account_created_at', 'verified', 'lang'

2 Data Pre-processing

The tweets specific to India are extracted from the dataset after removing all the unwanted attributes.
The resultant dataset contains 4 columns and 1574 rows.

	created_at	text	favourites_count	retweet_count
0	2020-03-29T00:04:06Z	"#Covid19: #SocialDistancing the Indian way" ...	10374	0
1	2020-03-29T00:04:55Z	Possible case of human to animal transmission ...	22880	0
2	2020-03-29T00:07:20Z	#Coronavirus: #US eclipses 120,000 confirmed c...	12968	6
3	2020-03-29T00:09:14Z	The bitter truth 😔\n#covid_19 #covid19india #I...	6	0
4	2020-03-29T00:13:20Z	#CoronavirusOutbreak All Churches shd realize...	38	0

3 Data Analysis

For Data Analysis all the stopwords, symbols, numbers and special characters are removed from the text attribute.
Now, the output of pre-processed tweets in the text attribute becomes more meaningful and readable when compared to the collected tweets.

```
0                  covid19 socialdistancing the indian way
1     possible case of human to animal transmission ...
2     coronavirus us eclipses 120 000 confirmed case...
3     the bitter truth covid 19 covid19india lockdow...
4     coronavirusoutbreak all churches shd realize t...
Name: text, dtype: object
```

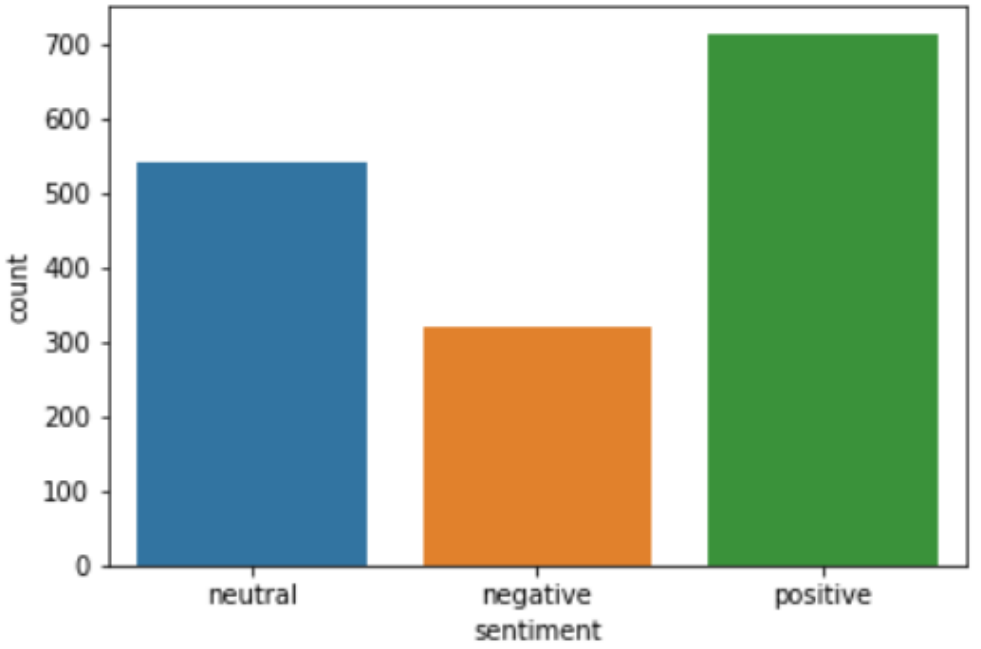
4 Sentiment Analysis

Sentiment analysis is being done using one of the NLP library 'TextBlob' under IBM Watson Notebook. We created a new dataset which stores sentiments (positive, negative or neutral) and polarity corresponding to all tweets.

	created_at	text	favourites_count	retweet_count	sentiment	polarity
0	0.066667	covid19 socialdistancing the indian way	10374	0	neutral	0
1	0.066667	possible case of human to animal transmission ...	22880	0	negative	-0.23125
2	0.116667	coronavirus us eclipses 120 000 confirmed case...	12968	6	positive	0.4
3	0.150000	the bitter truth covid 19 covid19india lockdow...	6	0	negative	-0.1
4	0.216667	coronavirusoutbreak all churches shd realize t...	38	0	positive	0.1375

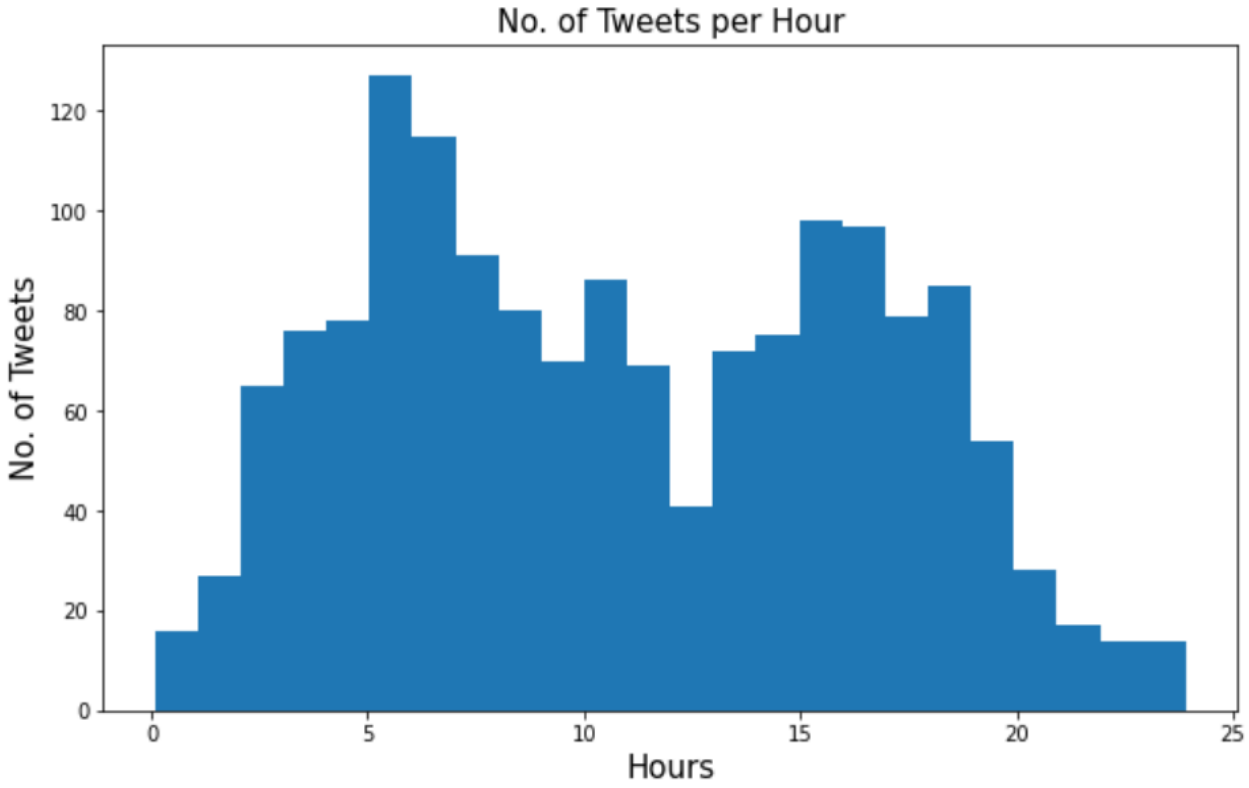
```
positive    715
neutral     540
negative    319
```

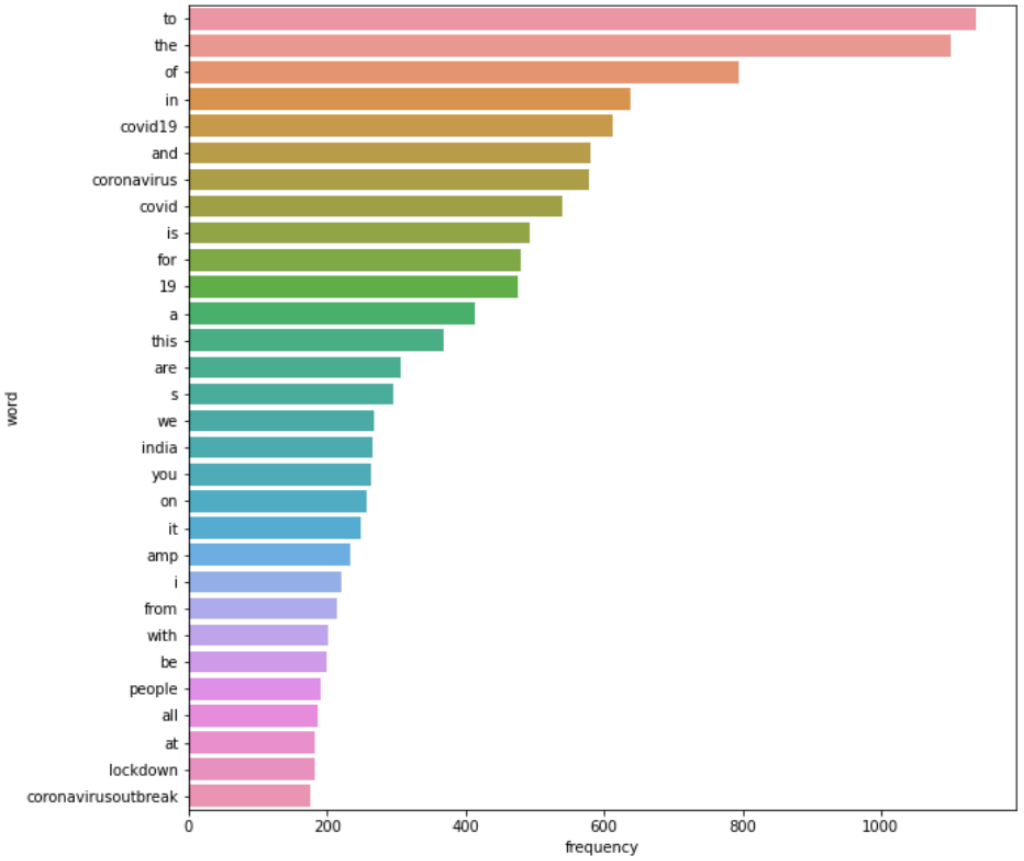
```
Name: sentiment, dtype: int64
```

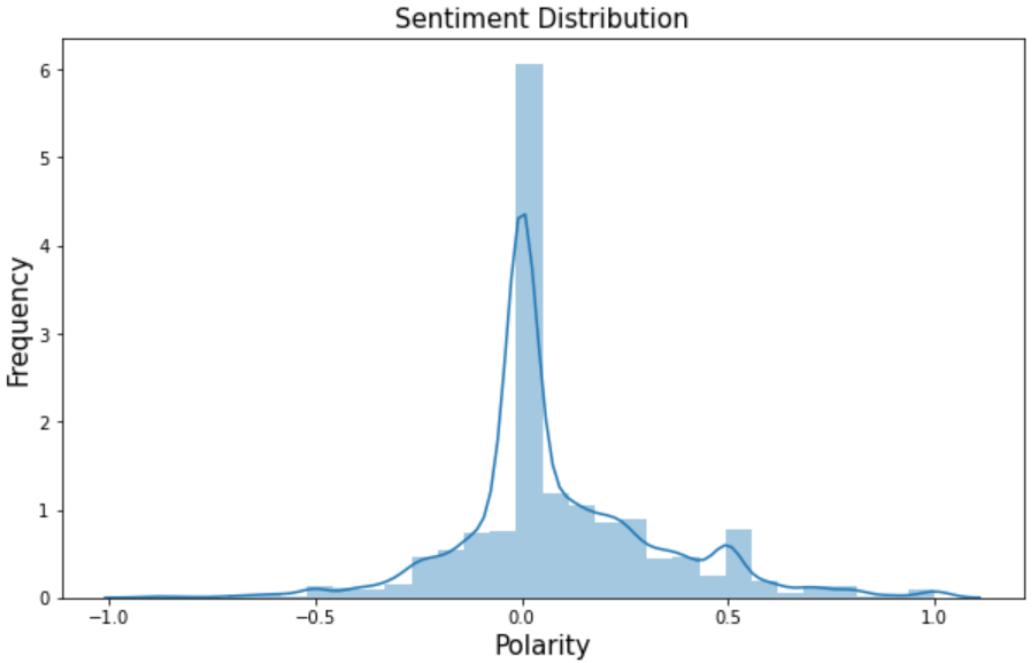


4 Visualization

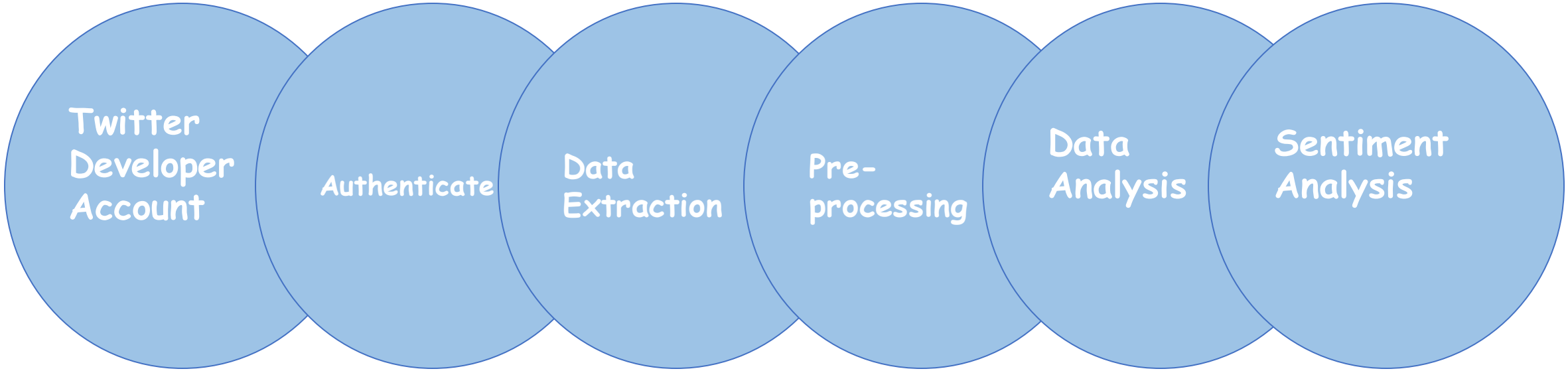
Many visualization plots from libraries like matplotlib and seaborn are used to visualize the results.




to
the
of
in
covid19
and
coronavirus
covid
is
for
19
a
this
are
s
we
india
you
on
it
amp
i
from
with
be
people
all
at
lockdown
coronavirusoutbreak
word
0
200
400
600
800
1000
frequency


Sentiment Distribution
Frequency
0
1
2
3
4
5
6
−1.0
−0.5
0.0
0.5
1.0
Polarity

5 Latest Tweets Retrieval

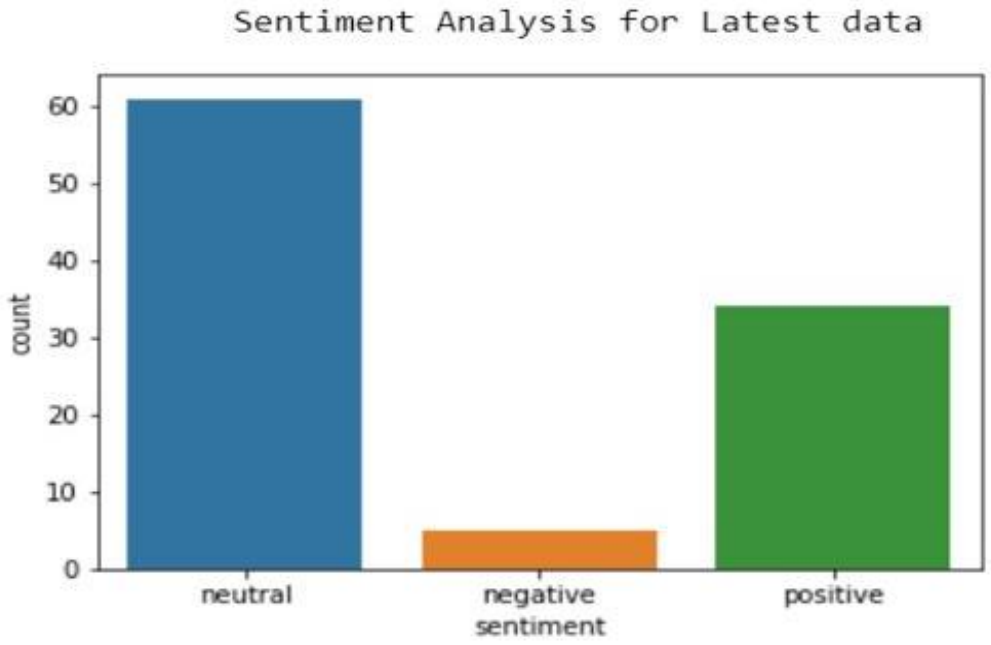


5 Latest Tweets Sentiment Analysis

Dataset with Sentiment Analysis values, i.e.,sentiment and polarity.

	Tweets	len	ID	Date	Source	Likes	RTs	sentiment	polarity
0	When will this lockdown end \n#COVID19 #lockdo...	86	1283076566767988736	2020-07-14 16:31:06	Twitter for Android	0	0	neutral	0
1	If #Deewar was filmed in 2020 😐 \n#lockdown #l...	81	1283071767586594816	2020-07-14 16:12:02	Twitter for Android	1	0	neutral	0
2	#कड़वासच #Bittertruth \n\With due credits to car...	111	1283068511263186944	2020-07-14 15:59:06	Twitter for Android	0	0	negative	-0.125
3	Potato Fenugreek Leaves Buns\n\n#lockdownindia...	137	1283060781047668736	2020-07-14 15:28:22	Twitter for Android	1	1	neutral	0
4	#thekingofkhansra #changez #changezandme #lock...	136	1283057041184100354	2020-07-14 15:13:31	Instagram	0	0	neutral	0

Name: sentiment, dtype: int64

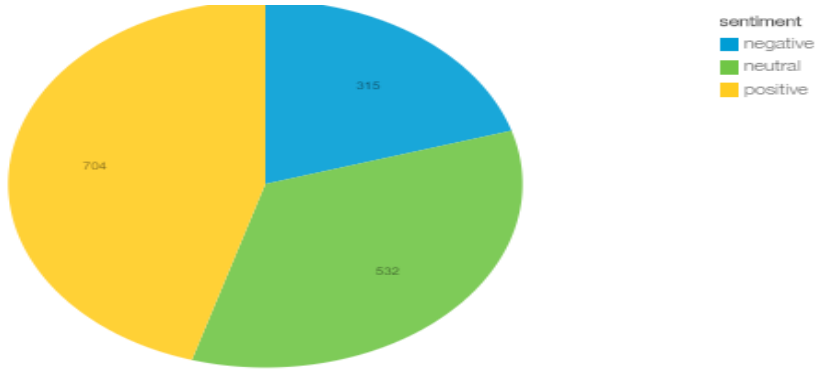


positive 715
neutral 540
negative 319

6 Dashboard Creation

Dashboard is imported in project using IBM Watson Studio with the analysed data and shows interactive UI to know the patterns, sentiments and polarity of tweets belonging to selected variation of sentiment.

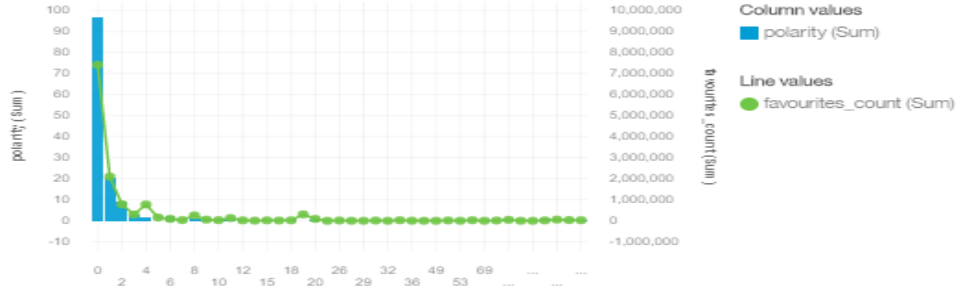
1 more storm in pune we felt like thunder jus hit our home der was spark in switch boards n above the
Summary
1 ve case 5 people then 5 to 25 and goes on coronavirushoax coronavirusoutbreak coronavirusindia
Summary
1000 covid 19 cases in as of today next few weeks are gonna be very very crucial indiafightcorona
Summary
13 lakh railway employees will donate rs 151 crore to pm cares fund to help combat coronavirus this e
Summary
15thmarch 2020 to 29thmarch 2020 the death ratio increased from 13 to 18 increased by 5 in just 14 c



negative

neutral

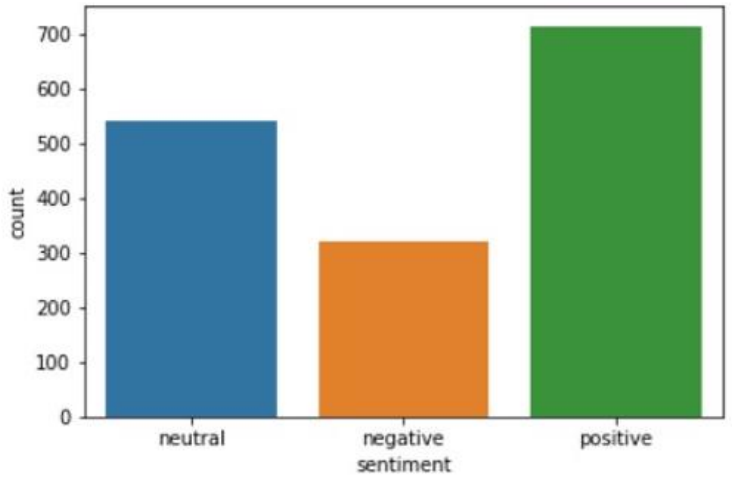
positive

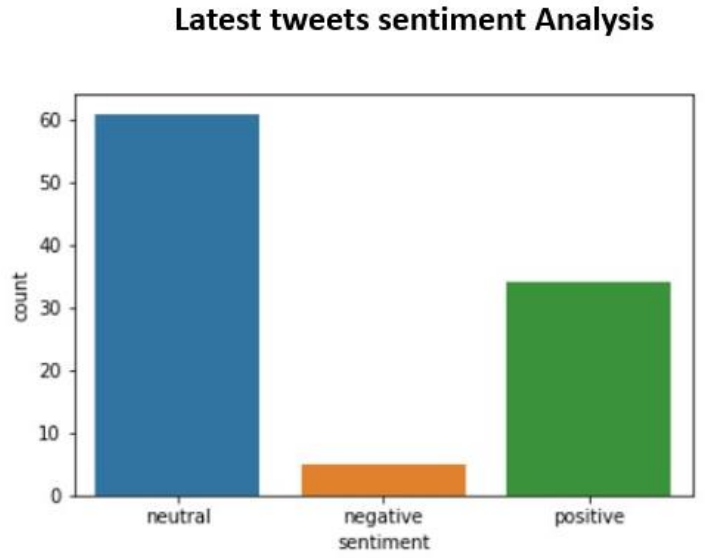


3,654 139.23

7 Results of Sentiment Analysis

- During the initial phase, the score of positive reactions is more than that of the other sentiments.
- Observations from the recent analysis show decrease in the positive sentiment score.





Thanks to

&