# CORONAVIRUS TWEET SENTIMENT

# Sudhanshu Chouhan,Kapil Narayan Singh

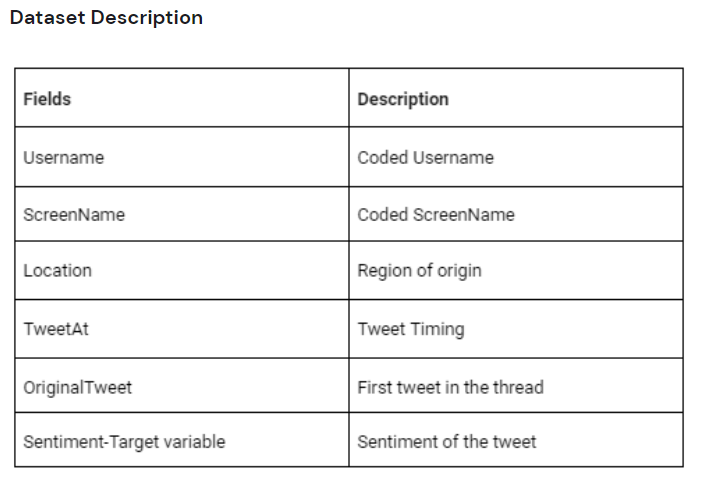
# Data scientist Trainee,Alma better

Abstract:

In this project we are going to discuss the various aspects of mantle aspects during the coronavirus pandemic. This once in an era pandemic teaches us various lessons and taking care of yourself and the people of the globe has now become the concern of every government policy manifesto.In a world full of cut throat competition to earn bread and butter we forget to build a good lifestyle and give less focus to our mantle health. This pandemic turns out to be a blessing in disguise for us and we start giving importance to our mantle health by doing yoga, meditation and what not to rejuvenate ourselves.

1. Problem Statement

This challenge will ask us to build a classification model to predict the sentiment of COVID-19 tweets. The tweets have been pulled from Twitter and manual tagging has been done then. The names and Usernames have been given in codes to avoid any privacy concern.

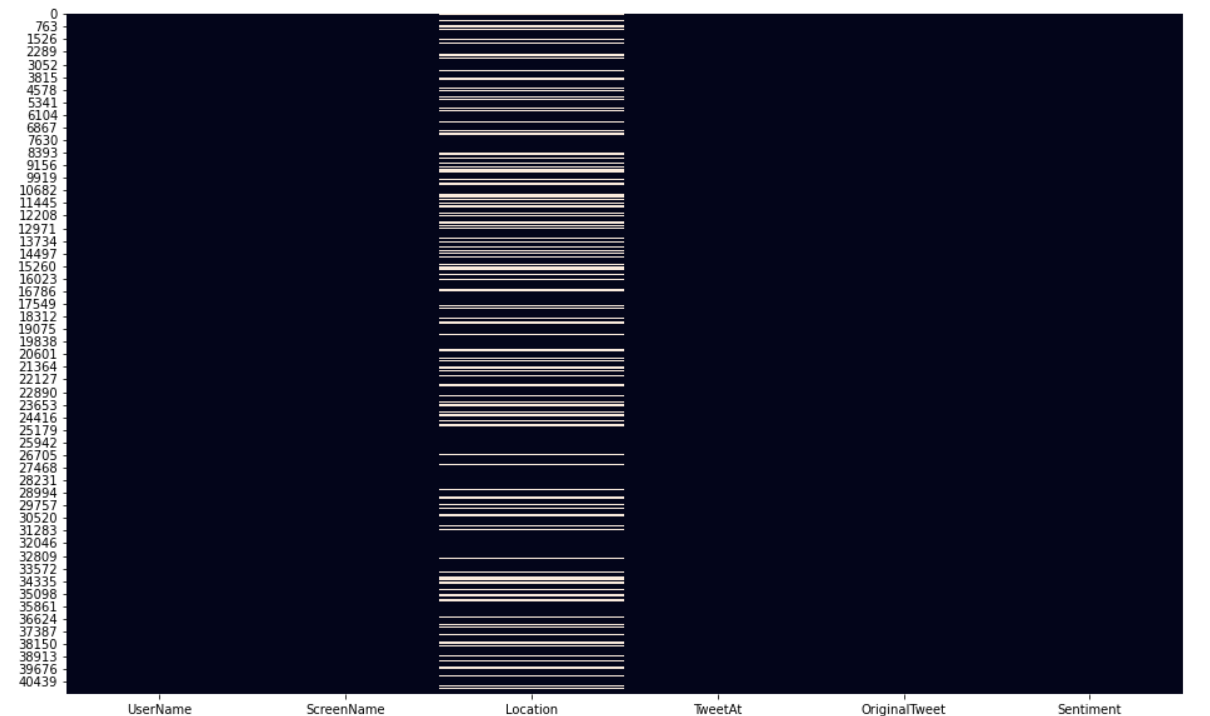


Introduction:

Exploratory Data Analysis:

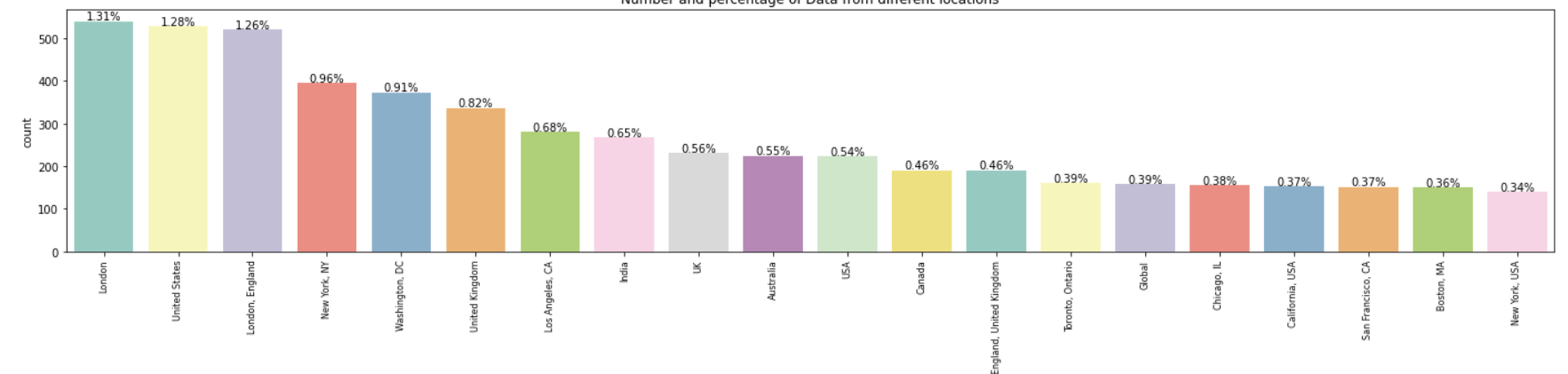
1.**Null values.**

We have performed analysis on whether there is presence of null values or not in our data.

****

Then we get to know that only location column has null values . So, for analysis purpose we have removed that column.

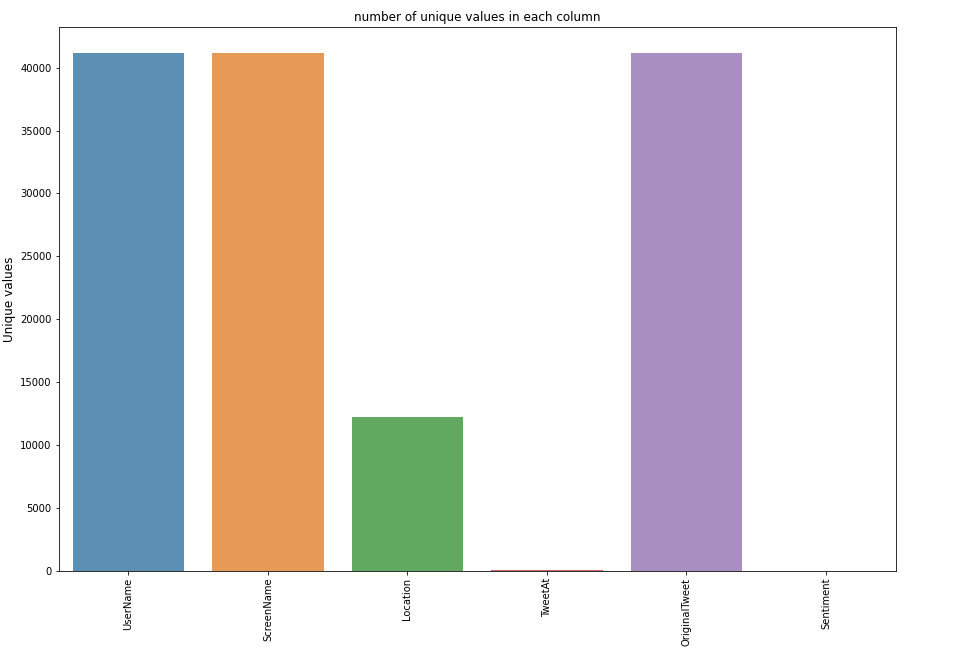
**2.On the basis of countries.**

****

On the Basis of countries we get to know that the U.K ,U.S.A, India and Canada are the countries from which a large number of tweets have been made.

**3.Number of unique values.**

From this analysis we get to know the unique values in each column and witnessed that our target column has 5 different type of outputs on which we have to work and predict the sentiment of tweets that are made.



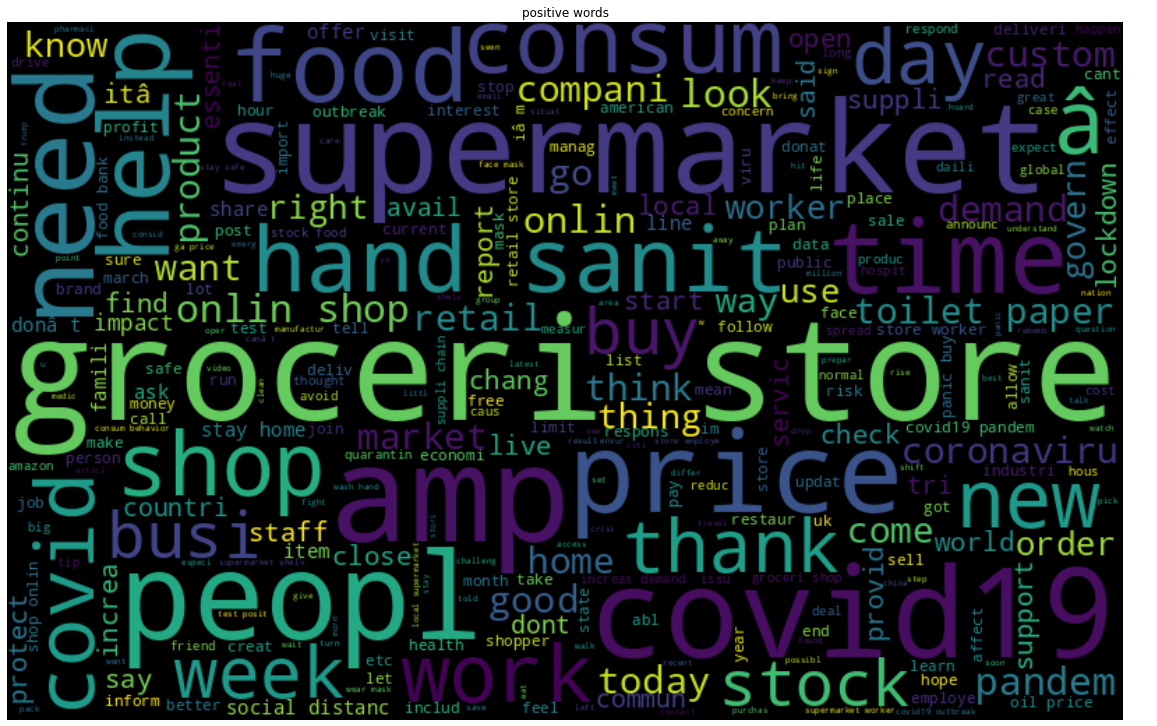
**4. Wordcloud**

In this analysis we get to know which type of word is most frequent and how many times that they have been used during the pandemic while tweeting.

****

**5.Positive words.**

We can see that there is not much difference between the Positive words and extremely positive words.

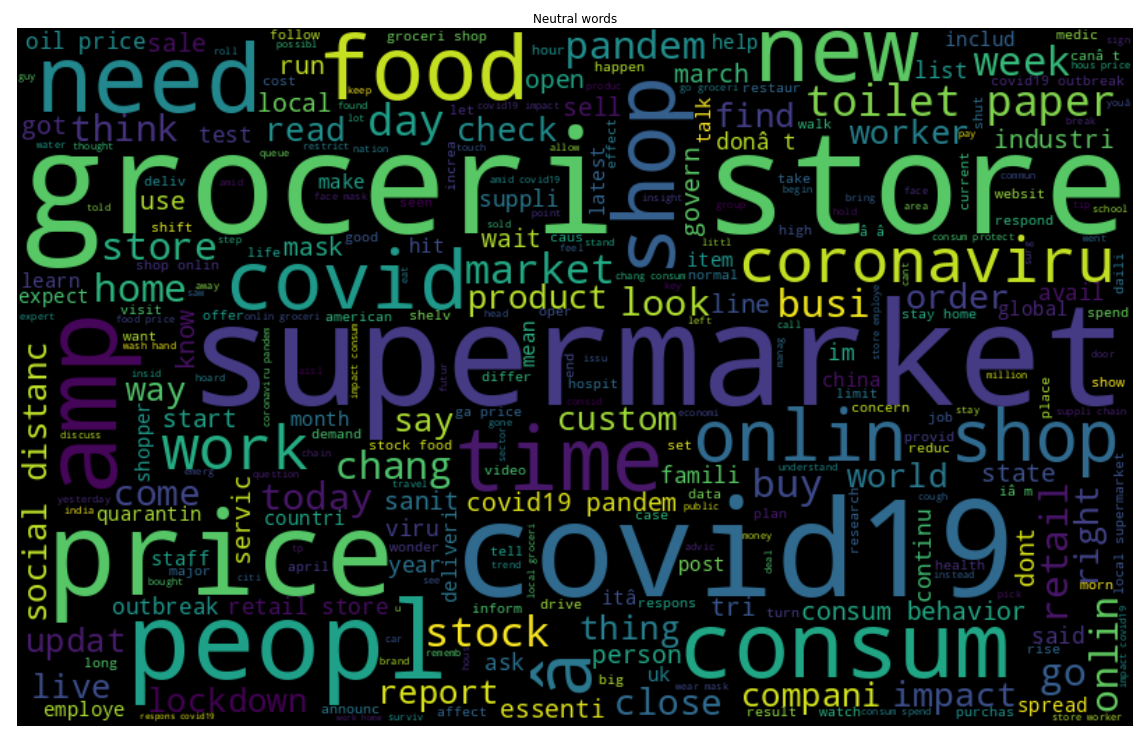
****

**6.Extremely Positive words**

Here we can see that Supermarket people, time work and sanitizer are the most used words in the extremely positive type of statement.****

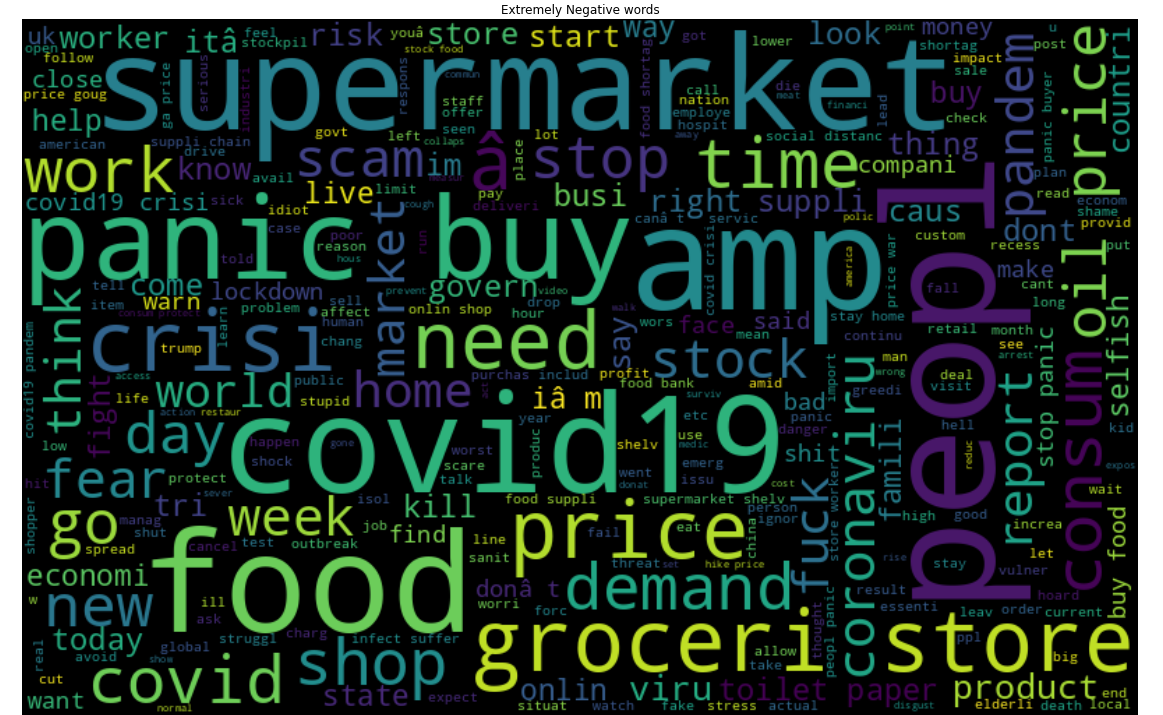
**7.Neutral words**

Neutral words follows as positive type of word composition and there is no such difference.

****

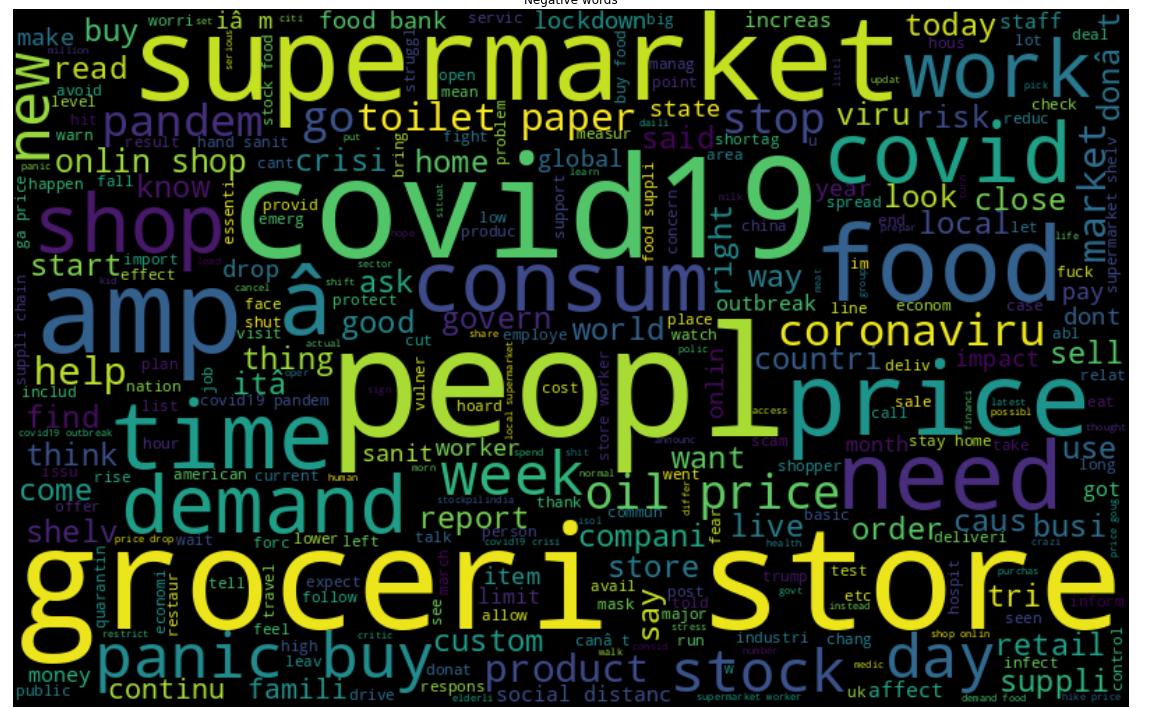
**8.Extremely negative word**

From this analysis also we get to know that the same type of word has been used. There is no such difference that can be seen but the usage of negative word have increased like price, demand, crisis, oil prices and panic.

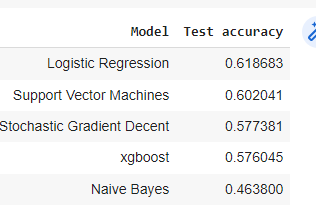
****

**9.Negative words**

Here, like extremely positive and positive tweets an analogy has been developed the same analogy has been seen in negative types of tweets.

****

* **Algorithm for multi classification**

****

From this analysis we get to know that logistic classification is the best classification if we are talking about multi classification.

* **Algorithm for binary classification**

****

After doing multi classification we have done binary classification and get to know that Stochastic Gradient Descent is the best performing algorithm and thus we opted this algorithm .

**Conclusion**

From the above discussion we first evaluated the model on the basis of multiple classification where we have taken multiple sentiments into consideration which includes positive ,negative,neutral, extremely positive and extremely negative sentiment.

* After applying various classification algorithms we come to a conclusion that the best is stochastic gradient Descent with a score of 86.582 percent.
* After multiple classification we divided our data into binary classification, in which extremely positive, positive and neutral are taken as 0 and extremely negative and negative are taken as 1 for classification.
* The output after Deploying various Models the best algorithm comes out to be Stochastic Gradient Descent.

**REFERENCES**

* NUMPY, PANDAS, MATPLOTLIB,SEABORN DOCUMENTATION.
* ALMA BETTER CLASS.
* ALMABETTER STUDENT COMMUNITY.