

# Project

**Topic :-**

**Data Analysis on how online shopping influences our Transportation Services.**

**Report by :-**

**Ravi Kumar Sinha**

B.Tech at Indian Institute of Technology  
(Banaras Hindu University), Varanasi

## Twitter Sentiment Analysis

### Sentiment Analysis:-

Sentiment analysis is the automated process of identifying and classifying subjective information in text data. This might be an opinion, a judgment, or a feeling about a particular topic or product feature.

The most common type of sentiment analysis is 'polarity detection' and involves classifying statements as Positive, Negative or Neutral.

Sentiment analysis uses Natural Language Processing (NLP) to make sense of human language and machine learning language to automatically deliver accurate results.

### Tools

In this project I have used following listed tools\libraries:-

#### snsrape :-

To extract data(tweets) from the Twitter I have used snsrape library.

Basically, snsrape is a scraping tool for social networking services (SNS). It scrapes things like users, user profiles, hashtags, searches, threads, list posts and returns the discovered items without using Twitter's API.

#### NumPy:-

I have imported NumPy for mathematical and logistics operations.

## **Pandas: -**

I have imported pandas for many data analysis operations some of them like: -  
to make Data Frame from the discrete Tweets and to do sentimental analysis.

## **Matplotlib :-**

I have imported matplotlib through which I able to data visualisation as a form of 'scatter plot,' 'Bar Graph' and 'Pie chart' and many comparisons among various results.

## **NLTK :-**

The Natural Language Toolkit (NLTK) is a platform used for building Python programs that work with human language data for applying in statistical natural language processing (NLP)

**Topic: -**

Data Analysis on how online shopping influences our Transportation Services.

**Description: -**

As we know the exponential growth of online shopping increasing day by day.

Many people want to buy the product online as of they get wide selection and variety of products with easy mode of payment most importantly, they also get "Fast Delivery Services and Replacement Services".

As you notice for betterment and satisfaction of the customer all the Online Store needs a well-developed "Transportation Services System".

**Analysis: -**

Here I am going to do Sentimental Analysis on "Online Shopping" as well as "Delivery Services System" under the following listed parameters: -

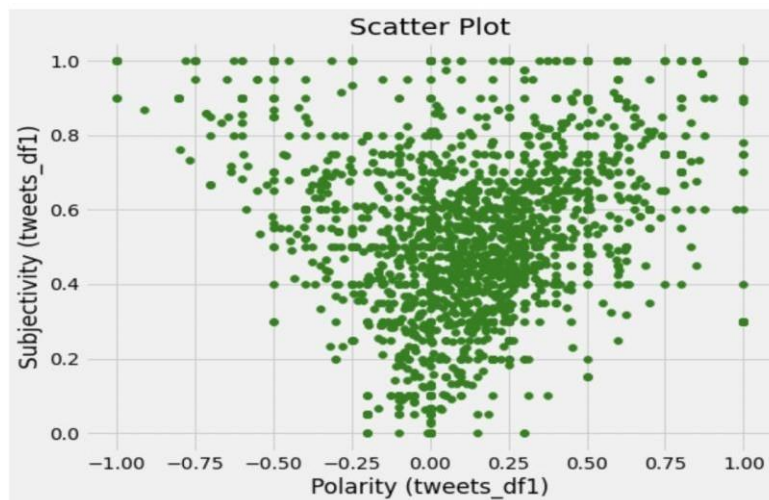
1. Sentimental Analysis on online shopping before covid pandemic.
2. Sentimental Analysis on online shopping during covid pandemic.
3. Sentimental Analysis on willingness of adopting online shopping.
4. Sentimental analysis on delivery Services System in online shopping.

## (1). Sentimental Analysis on online shopping before covid pandemic.

Here I have scrapped the tweets from the Twitter on the topic online shopping before covid pandemic between the timeline 2018-01-01 to 2019-12-31. Afterwards I cleaned the tweets by removing (@#\$%^&\*<>|\+/) and then assigned "Subjectivity" and "Polarity" to the tweets based on some algorithms.

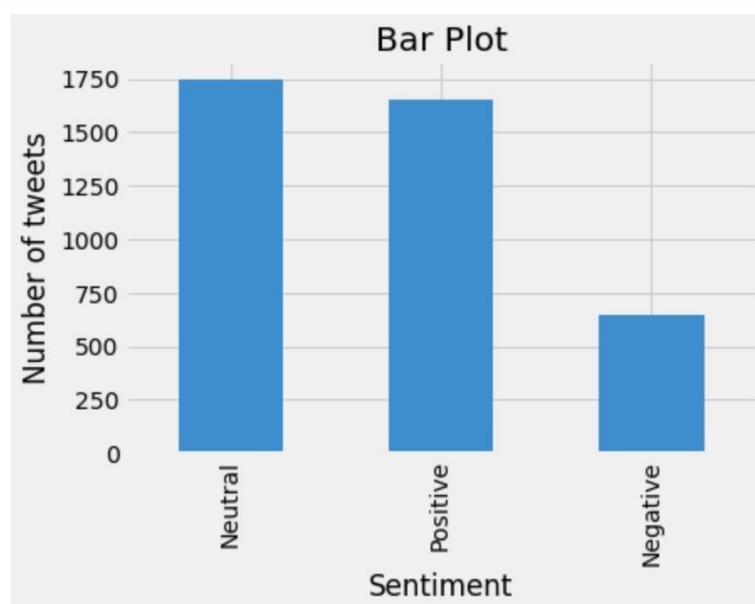
### tweets\_list1 : Scatter Plot : Subjectivity VS Polarity

This shows the relation how the polarity changes accordingly to subjectivity.



### tweets\_list1 : Sentiment: Bar Graph

This shows the pictorial representation of "Number of Tweets" comes under the "Positive", "Negative" and "Neural" sentiment.



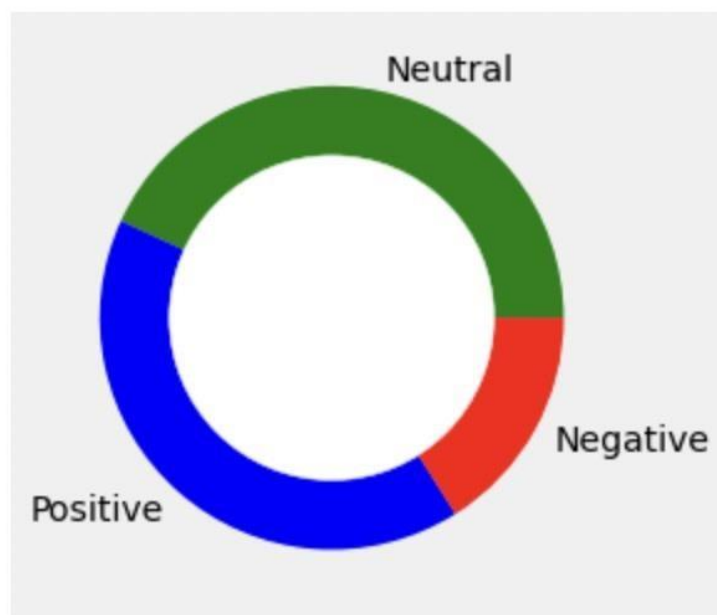
tweets\_list1 : tweet\_count : percentage

This shows the actual "Number of Tweets" and "Percentage" comes under the "Positive" , "Negative" and "Neural" sentiment.

	Total	Percentage
Neutral	1746	43.14
Positive	1652	40.82
Negative	649	16.04

tweets\_list1 : tweet\_count : pie chart

This shows the pictorial representation of the actual "Number of Tweets" comes under the "Positive" , "Negative" and "Neural" sentiment.



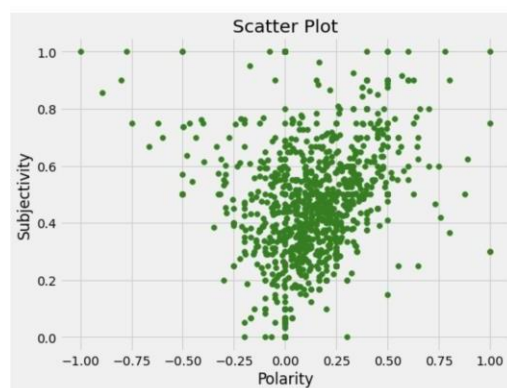
## (2). Sentimental Analysis on online shopping during covid pandemic.

Here I have scrapped the tweets from the Twitter on the topic online shopping before covid pandemic between the timeline 2020-01-01 to 2022-01-31.

Afterwards I cleaned the tweets by removing (@#\$%^&\*<>|\\+/) and then assigned "Subjectivity" and "Polarity" to the tweets based on some algorithms.

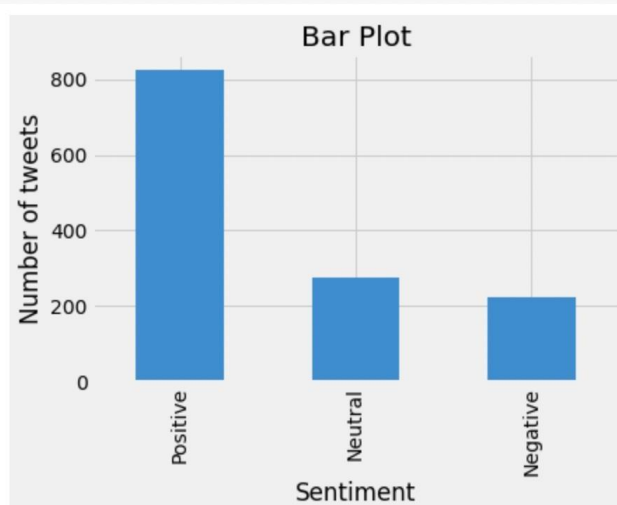
### tweets\_list2 : Scatter Plot : Subjectivity VS Polarity

This shows the relation how the polarity changes accordingly to subjectivity.



### tweets\_list2 : Sentiment: Bar Graph

This shows the pictorial representation of "Number of Tweets" comes under the "Positive", "Negative" and "Neural" sentiment.



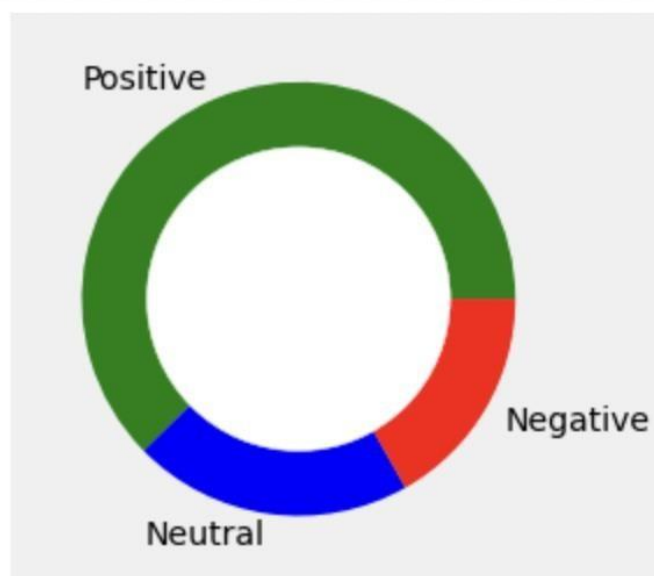
**tweets\_list2 : tweet\_count : percentage**

This shows the actual "Number of Tweets" and "Percentage" comes under the "Positive" , "Negative" and "Neural" sentiment.

	<b>Total</b>	<b>Percentage</b>
<b>Positive</b>	824	62.33
<b>Neutral</b>	276	20.88
<b>Negative</b>	222	16.79

**tweets\_list12: tweet\_count : pie chart**

This shows the pictorial representation of the actual "Number of Tweets" comes under the "Positive" , "Negative" and "Neural" sentiment.





### (3). Sentimental Analysis on willingness of adopting online shopping.

Here I have scrapped the tweets from the Twitter on the topic online

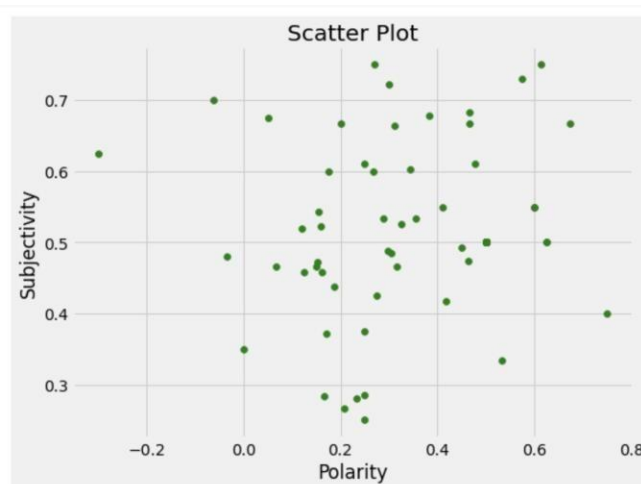
shopping before covid pandemic between the timeline :2018-01-01 to 2022-05-01.

Afterwards I cleaned the tweets by removing (@#\$\$%^&\*<>|\\+/-) and

then assigned "Subjectivity" and "Polarity" to the tweets based on some algorithms.

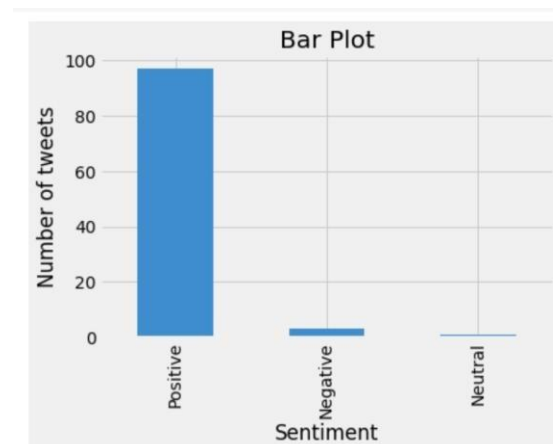
#### **tweets\_list3 : Scatter Plot : Subjectivity VS Polarity**

This shows the relation how the polarity changes accordingly to subjectivity.



#### **tweets\_list3: Sentiment: Bar Graph**

This shows the pictorial representation of "Number of Tweets" comes under the "Positive", "Negative" and "Neural" sentiment.



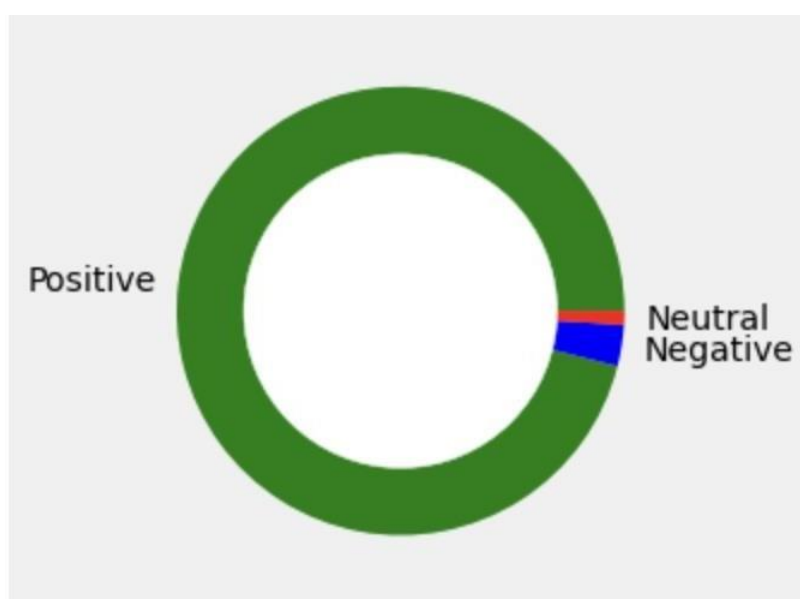
### tweets\_list3: tweet\_count : percentage

This shows the actual "Number of Tweets" and "Percentage" comes under the "Positive" , "Negative" and "Neural" sentiment.

	Total	Percentage
Positive	97	96.04
Negative	3	2.97
Neutral	1	0.99

### tweets\_list3: tweet\_count : pie chart

This shows the pictorial representation of the actual "Number of Tweets" comes under the "Positive", "Negative" and "Neural" sentiment.

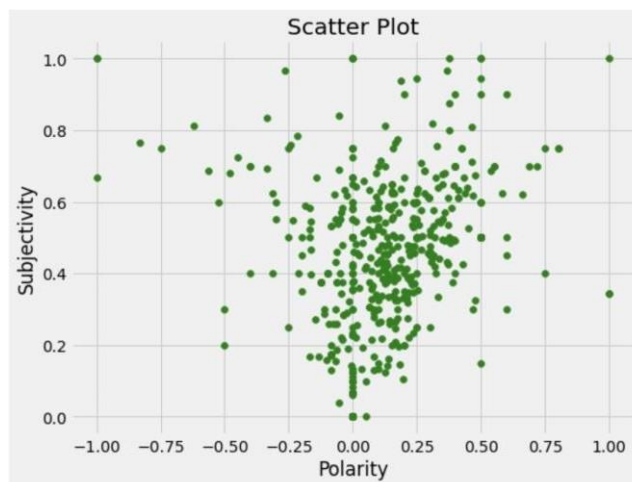


#### (4). Sentimental analysis on delivery Services System in online shopping

Here I have scrapped the tweets from the Twitter on the topic online shopping before covid pandemic between the timeline 2018-01-01 to 2022-05-01 Afterwards I cleaned the tweets by removing (@#\$%^&\*<>| \+/) and then assigned "Subjectivity" and "Polarity" to the tweets based on some algorithms.

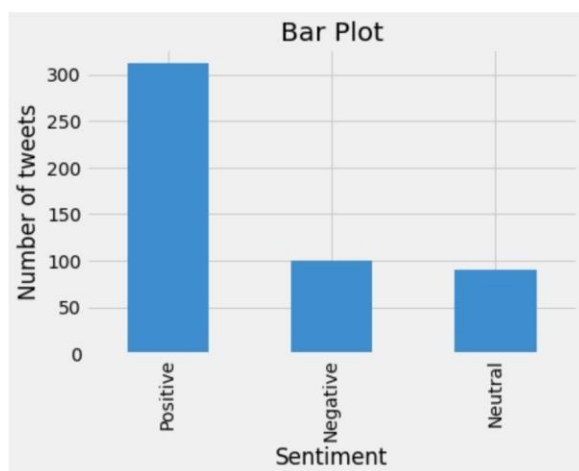
##### tweets\_list4: Scatter Plot: Subjectivity VS Polarity

This shows the relation how the polarity changes accordingly to subjectivity.



##### tweets\_list4: Sentiment: Bar Graph

This shows the pictorial representation of "Number of Tweets" comes under the "Positive", "Negative" and "Neutral" sentiment.



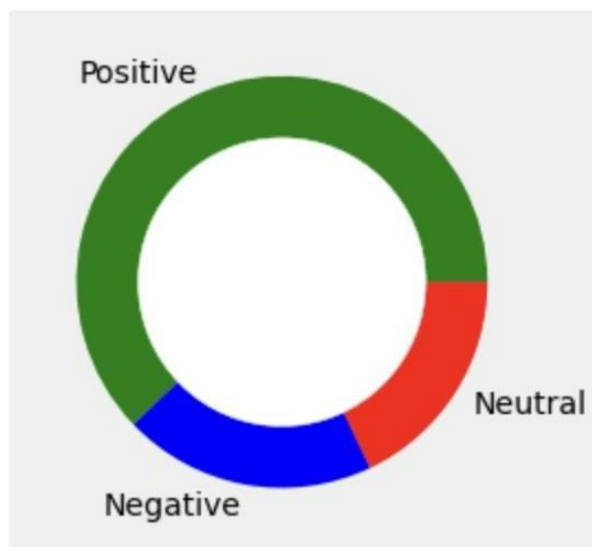
**tweets\_list4 : tweet\_count : percentage**

This shows the actual "Number of Tweets" and "Percentage" comes under the "Positive", "Negative" and "Neural" sentiment.

	Total	Percentage
Positive	312	62.15
Negative	100	19.92
Neutral	90	17.93

**tweets\_list4 : tweet\_count : pie chart**

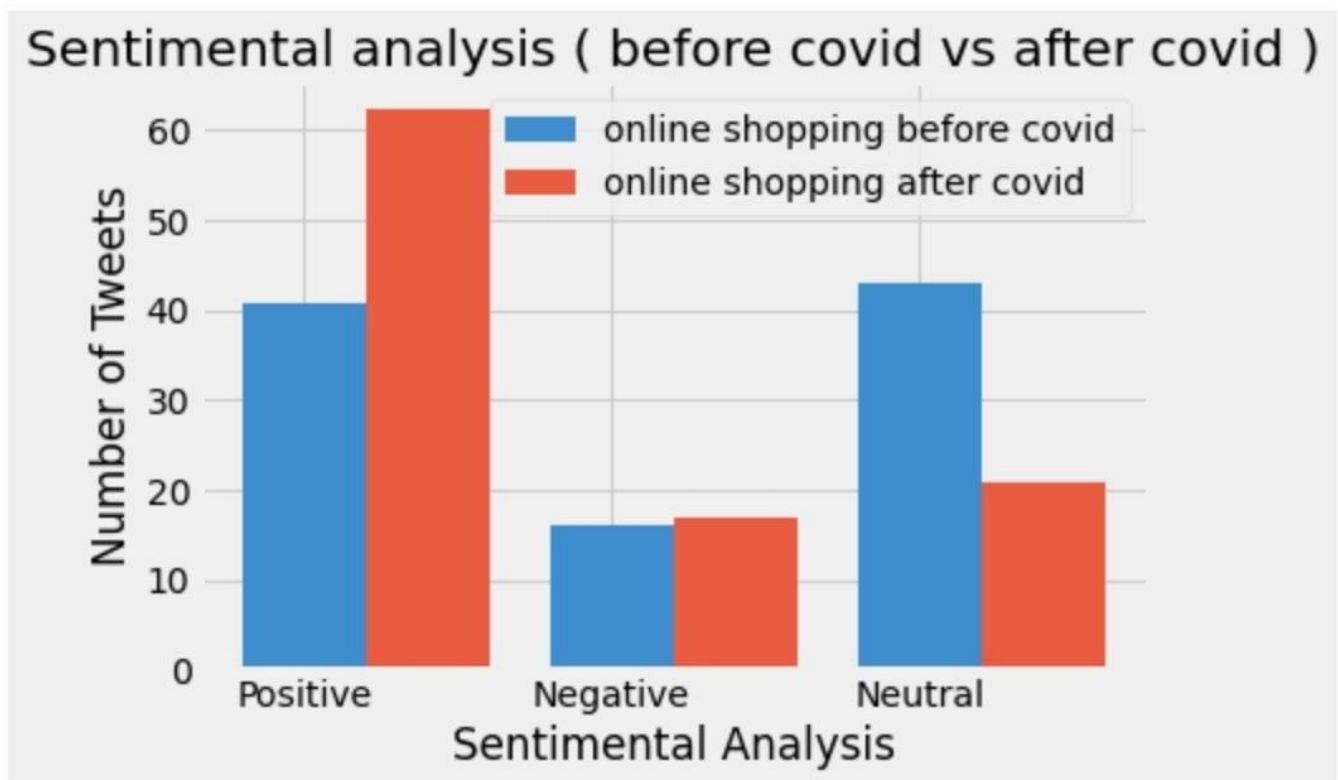
This shows the pictorial representation of the actual "Number of Tweets" comes under the "Positive", "Negative" and "Neural" sentiment.



## Comparison of Sentimental Analysis

(1). Comparison of Sentimental analysis on online shopping between before covid pandemic verses during covid pandemic.

Here the results show how the trend of online shopping influences by the covid pandemic.



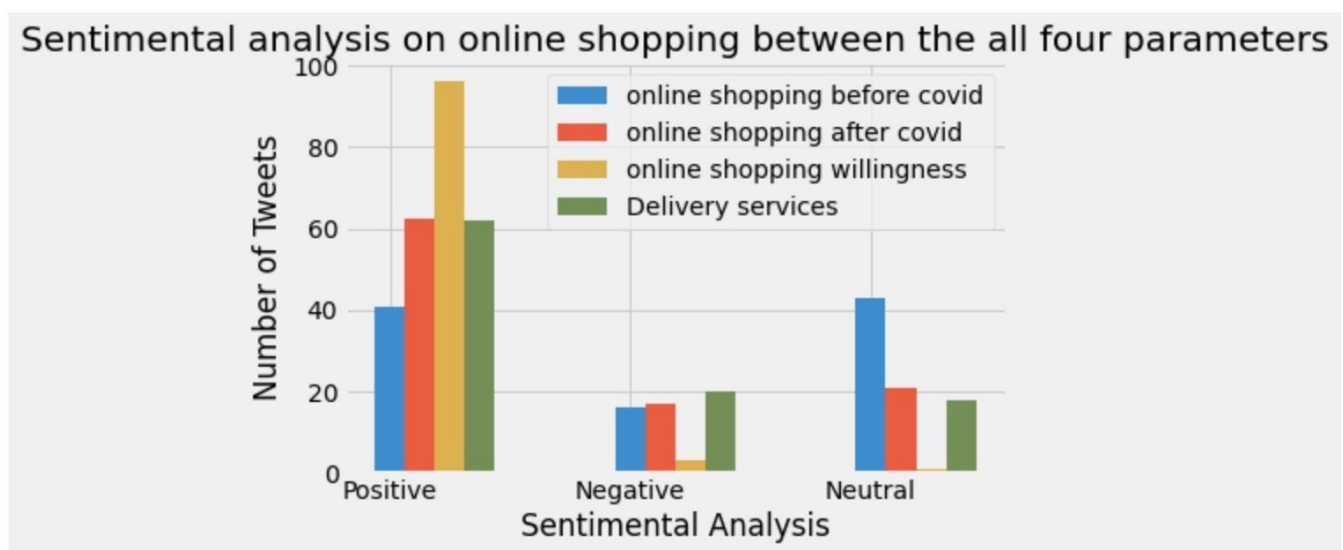
## (2). Comparison of Sentimental analysis on online shopping Willingness verses Delivery Services.

Here the results show how willingness of the people adopting the way of online shopping get influenced by our delivery services. (Transportation Services System)



**(3). Comparison of Sentimental analysis on online shopping between the all four parameters.**

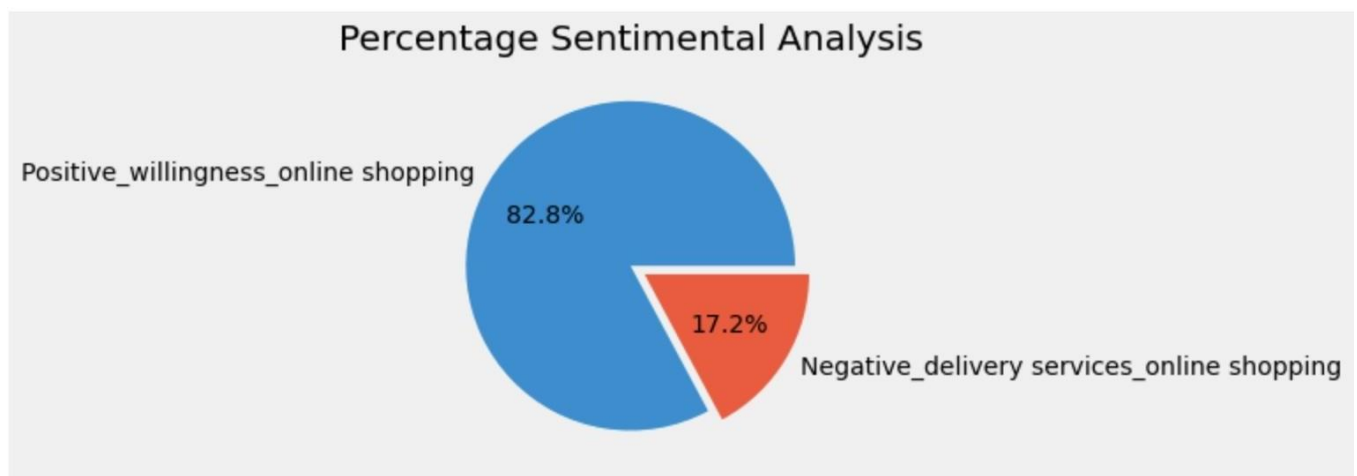
Here the results show the pictorial relation between all the four parameters.



## "Percentage Sentimental Analysis" on 'Positive\_willingness\_online shopping' verses 'Negative\_delivery services\_online shopping'

Here the result shows the relation between the percentage of people who wants to go for online shopping verses the percentage of people who faces the problems by the delivery Services.

So, as we know in upcoming years the trend and the growth of online shopping / e-marketing increases so we have to improve our "Delivery Services" or we can say "Transportation Services System" to eliminate the 17.2% of hindrance.



----- Thank You-----