

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import re
from matplotlib import style
style.use('ggplot')
from textblob import TextBlob
from nltk.tokenize import word_tokenize
from nltk.stem import PorterStemmer
from nltk.corpus import stopwords
stop_words=set(stopwords.words('english'))
from wordcloud import WordCloud
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix,classification_report
```

```
In [2]: df=pd.read_csv('C:\\\\Users\\\\LENOVO\\\\Documents\\\\peter_obi_august.csv')
```

```
In [3]: df.head(6)
```

Out[3]:

	Datetime	Tweet_Id	Tweets	Username	Likes_count	Retweet_c
0	2022-08-04 03:26:34+00:00	1555032419572531201	Look at this moron...What country are you taking...	lawalt2010	0	
1	2022-08-04 03:25:51+00:00	1555032236889620480	I feel Peter Obi and his team should as a matt...	EAGeorge01	0	
2	2022-08-04 03:24:45+00:00	1555031960564559873	I'm going to @JeffreyGuterman's upcoming Space...	DoomsBabe	0	
3	2022-08-04 03:24:26+00:00	1555031882135478279	True talk bro. That's the kind of people we never...	UtaziAlexander	0	
4	2022-08-04 03:24:07+00:00	1555031803857182721	Peter Obi already told us we've a Spending problem...	Fullness_101	0	
5	2022-08-04 03:16:14+00:00	1555029817430532103	Why are the APC people trying hard to disassociate...	MercyEgbai	2	

```
In [4]: df.shape
```

Out[4]: (18883, 10)

```
In [5]: df.isnull().sum()
```

```
Out[5]: Datetime      0
Tweet_Id       0
Tweets         0
Username       0
Likes_count    0
Retweet_count  0
Quotes_count   0
Replies_count  0
Followers_count 0
Source          0
dtype: int64
```

```
In [6]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18883 entries, 0 to 18882
Data columns (total 10 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Datetime        18883 non-null   object  
 1   Tweet_Id        18883 non-null   int64  
 2   Tweets          18883 non-null   object  
 3   Username        18883 non-null   object  
 4   Likes_count     18883 non-null   int64  
 5   Retweet_count   18883 non-null   int64  
 6   Quotes_count    18883 non-null   int64  
 7   Replies_count   18883 non-null   int64  
 8   Followers_count 18883 non-null   int64  
 9   Source          18883 non-null   object  
dtypes: int64(6), object(4)
memory usage: 1.4+ MB
```

```
In [7]: df.columns
```

```
Out[7]: Index(['Datetime', 'Tweet_Id', 'Tweets', 'Username', 'Likes_count',
               'Retweet_count', 'Quotes_count', 'Replies_count', 'Followers_count',
               'Source'],
              dtype='object')
```

```
In [8]: df.drop(['Datetime', 'Tweet_Id', 'Username', 'Likes_count',
               'Retweet_count', 'Quotes_count', 'Replies_count', 'Followers_count',
               'Source'], axis=1, inplace=True)
```

In [9]: df.head(5)

Out[9]:

	Tweets
0	Look at this moron...What country are you taki...
1	I feel Peter Obi and his team should as a matt...
2	I'm going to @JeffreyGuterman's upcoming Space...
3	True talk bro. That's the kind of people we ne...
4	Peter Obi already told us we've a Spending pro...

In [10]:

```
print(df['Tweets'].iloc[0], "\n")
print(df['Tweets'].iloc[1], "\n")
print(df['Tweets'].iloc[2], "\n")
print(df['Tweets'].iloc[3], "\n")
print(df['Tweets'].iloc[4], "\n")
```

Look at this moron...What country are you taking over???...It is an obituary of common sense for anyone to think Peter Obi can ever win 2023 presidential election...Sam Omatseye was right, Obidients are terrorists and Peter Obi is their commander...

I feel Peter Obi and his team should as a matter of urgency report this to the relevant authorities.

This outrageous article should not be handled with kid gloves.

Something is not right somewhere 😕

@PeterObi

@Morris_Monye

@doyinokupe

@NgLabour

@VictorIsrael_

@FS_Yusuf_ <https://t.co/gEf8UQXExD> (<https://t.co/gEf8UQXExD>)

I'm going to @JeffreyGuterman's upcoming Space. Will you join too? <https://t.co/RC2YTdnvIw> (<https://t.co/RC2YTdnvIw>)

True talk bro. That's the kind of people we need in this country. Peter Obi all the way.

Peter Obi already told us we've a Spending problem in this country. No one believed him. We do have a spending problem. How is our own security highly deficient but we find 1.4billion 2 fund & equip another country's? What did we do to deserve these corrupt leaders? #Obidatti2023 <https://t.co/siDiuaIKDo> (<https://t.co/siDiuaIKDo>)

```
In [11]: def df_prepocessing (tweet):
    tweet=tweet.lower()
    tweet=re.sub(r"https\s+|www\S+https\S+", '',tweet,flags=re.MULTILINE)
    tweet=re.sub(r'@\w+|\#', '',tweet)
    tweet=re.sub(r'[^w\s]', '',tweet)
    tweet_token=word_tokenize(tweet)
    filter_tweet=[w for w in tweet_token if not w in stop_words]
    return " ".join(filter_tweet)
```

```
In [12]: df["Tweets"] = df[ 'Tweets' ].apply(df_prepocessing)
```

```
In [13]: df.head(5)
```

Out[13]:

Tweets

0	look moronwhat country taking overit obituary ...
1	feel peter obi team matter urgency report rele...
2	im going jeffreygutermans upcoming space join ...
3	true talk bro thats kind people need country p...
4	peter obi already told us weve spending proble...

```
In [14]: stemmer=PorterStemmer()
def stemming(data):
    tweet=[stemmer.stem(word) for word in data]
    return data
```

```
In [15]: df[ 'Tweets' ] = df[ 'Tweets' ].apply(lambda x: stemming(x))
```

```
In [16]: df.head(5)
```

Out[16]:

Tweets

0	look moronwhat country taking overit obituary ...
1	feel peter obi team matter urgency report rele...
2	im going jeffreygutermans upcoming space join ...
3	true talk bro thats kind people need country p...
4	peter obi already told us weve spending proble...

```
In [17]: print(df['Tweets'].iloc[0], "\n")
print(df['Tweets'].iloc[1], "\n")
print(df['Tweets'].iloc[2], "\n")
print(df['Tweets'].iloc[3], "\n")
print(df['Tweets'].iloc[4], "\n")
```

look moronwhat country taking overit obituary common sense anyone think peter o
bi ever win 2023 presidential electionsam omatseye right obidients terrorists p
eter obi commander

feel peter obi team matter urgency report relevant authorities outrageous artic
le handled kid gloves something right somewhere peterobi morris_monye doyinokup
e nglabour victorisrael_fs_yusuf_ httpstcogef8uqxexd

im going jeffreygutermans upcoming space join httpstcorc2ytdnviw

true talk bro thats kind people need country peter obi way

peter obi already told us weve spending problem country one believed spending p
roblem security highly deficient find 14billion 2 fund amp equip another countr
ys deserve corrupt leaders obidatti2023 httpstcosidiuaikdo

```
In [18]: def polarity (Tweets):
    return TextBlob(Tweets).sentiment.polarity
```

```
In [19]: df['polarity']=df['Tweets'].apply(polarity)
```

```
In [20]: df.head(5)
```

Out[20]:

	Tweets	polarity
0	look moronwhat country taking overit obituary ...	0.261905
1	feel peter obi team matter urgency report rele...	-0.104762
2	im going jeffreygutermans upcoming space join ...	0.000000
3	true talk bro thats kind people need country p...	0.475000
4	peter obi already told us weve spending proble...	-0.450000

```
In [21]: def sentiment (label):
    if label < 0:
        return "Negative"
    elif label == 0:
        return "Neutral"
    elif label > 0:
        return "Positive"
```

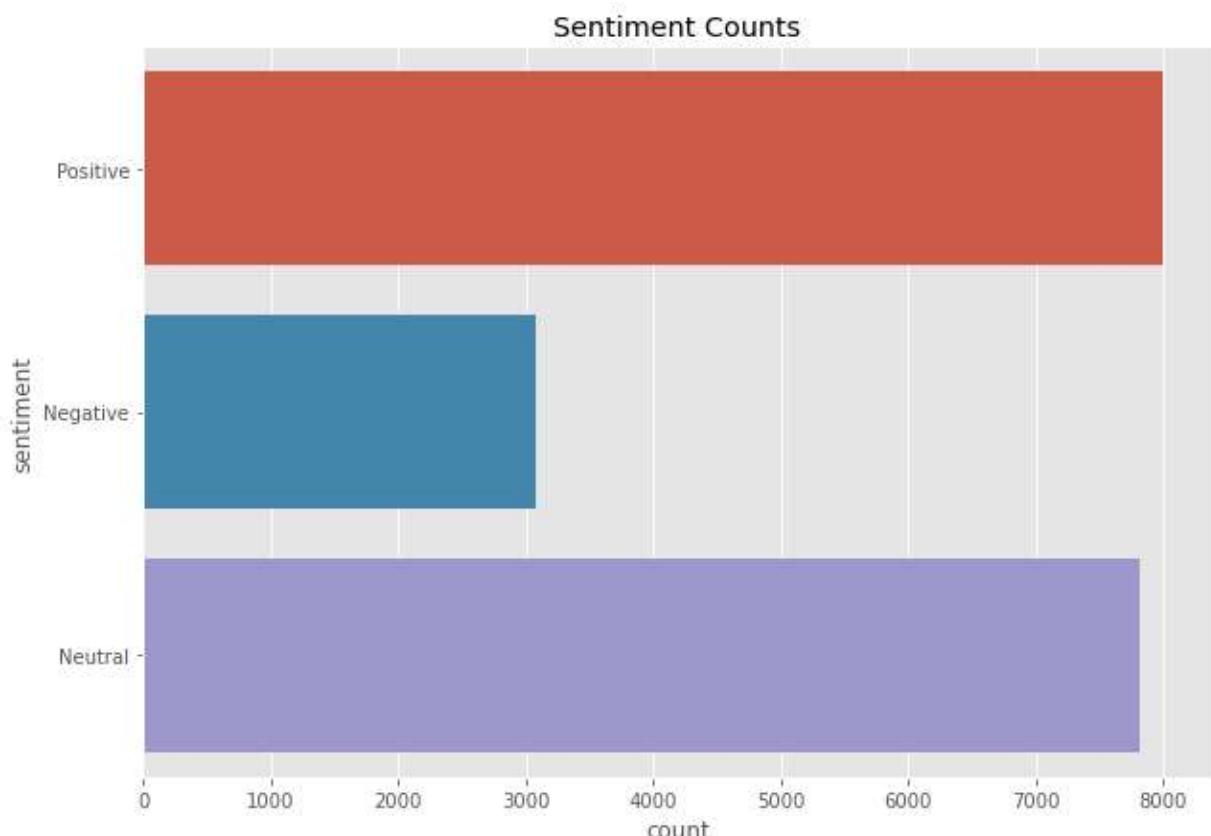
```
In [22]: df['sentiment']=df['polarity'].apply(sentiment)
```

In [23]: `df.head(5)`

Out[23]:

	Tweets	polarity	sentiment
0	look moronwhat country taking overit obituary ...	0.261905	Positive
1	feel peter obi team matter urgency report rele...	-0.104762	Negative
2	im going jeffreygutermans upcoming space join ...	0.000000	Neutral
3	true talk bro thats kind people need country p...	0.475000	Positive
4	peter obi already told us weve spending proble...	-0.450000	Negative

In [24]: `plt.figure(figsize=(10,7))
sns.countplot(y="sentiment", data=df).set_title('Sentiment Counts')
plt.show()`

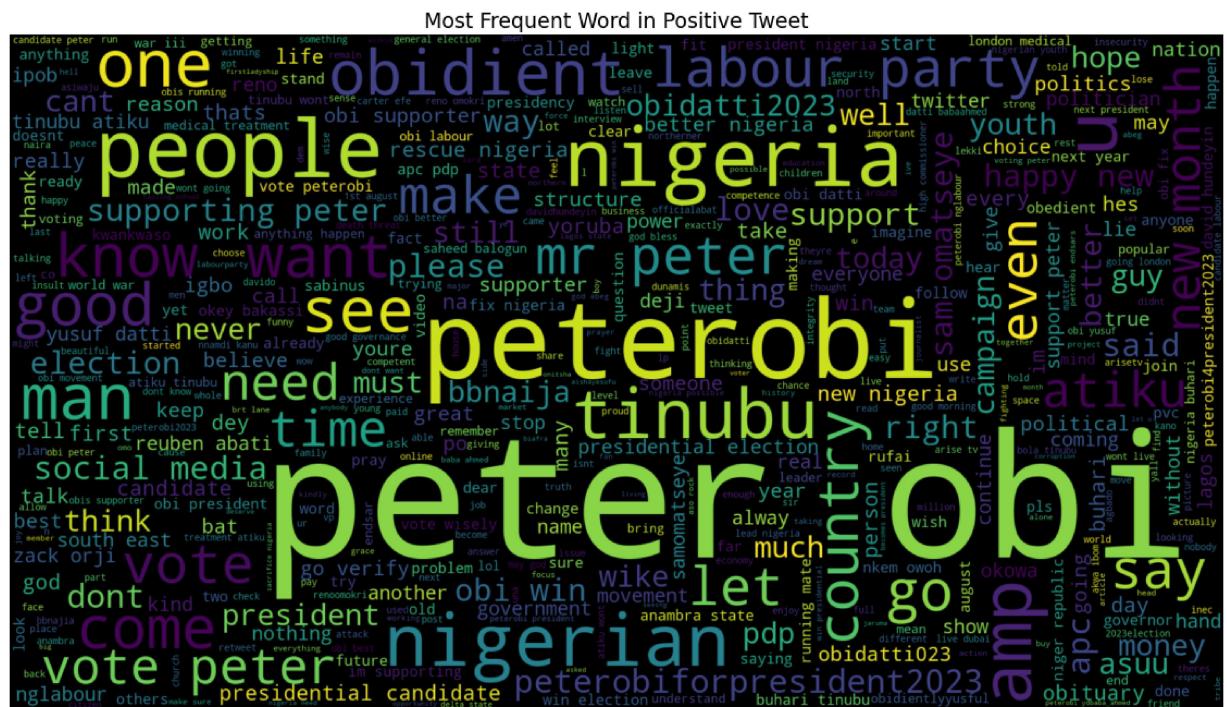


```
In [25]: pos_tweet=df[df.sentiment=="Positive"]
pos_tweet=pos_tweet.sort_values(['polarity'], ascending=False)
pos_tweet.head()
```

Out[25]:

	Tweets	polarity	sentiment
12164	peter obi messiah hes best candidate race	1.0	Positive
3640	best time buy offplan sabinus burna cardi pete...	1.0	Positive
2492	nigeria need give peterobi chance rule country...	1.0	Positive
8602	peter obi best agbado lokan httpstcozltlplo7c	1.0	Positive
2948	presidential aspirants peter obi best option	1.0	Positive

```
In [41]: tweet = ' '.join([word for word in pos_tweet["Tweets"]])
plt.figure(figsize=(20,14),facecolor='None')
wordcloud=WordCloud(max_words=500,width=1800,height=1000).generate(tweet)
plt.imshow(wordcloud,interpolation='bilinear')
plt.axis('off')
plt.title('Most Frequent Word in Positive Tweet',fontsize=19)
plt.savefig("Cloud.png1", format="png")
plt.show()
```

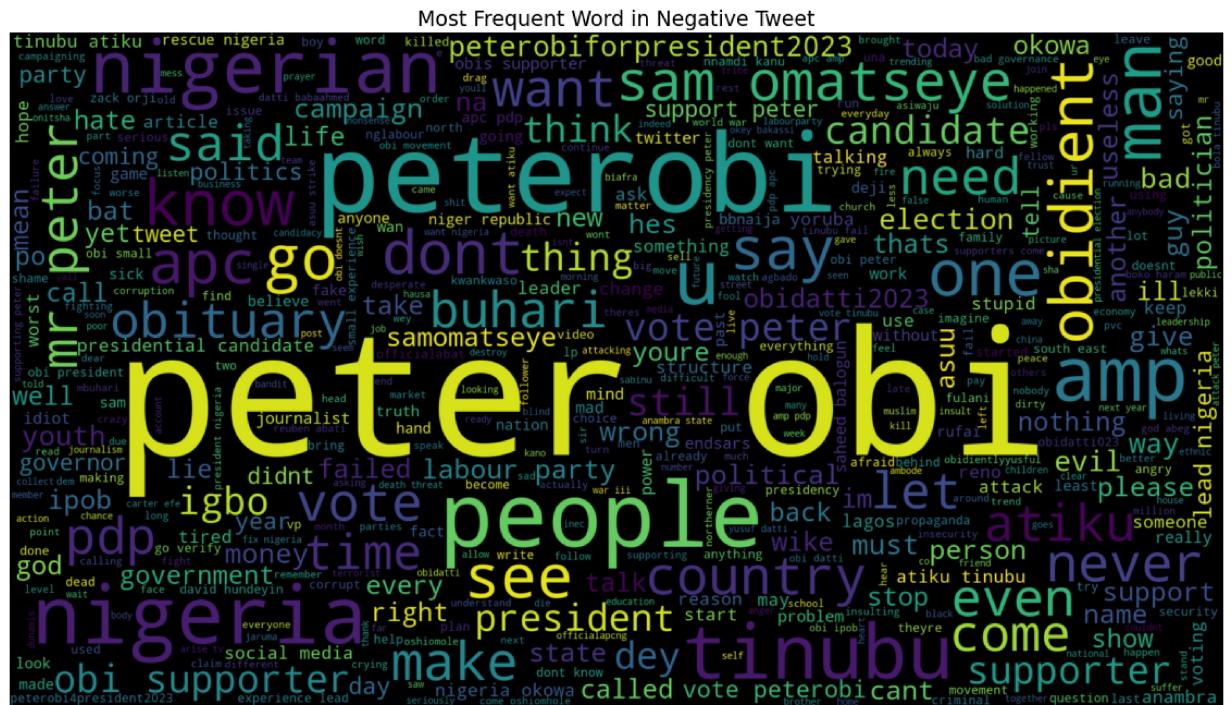


```
In [27]: neg_tweet=df[df.sentiment=="Negative"]
          neg_tweet=neg_tweet.sort_values(['polarity'], ascending=False)
          neg_tweet.head()
```

Out[27]:

		Tweets	polarity	sentiment
15261	people often unfairly amplify reaction obi sup...	-9.251859e-18		Negative
8356	wow old fool insulting peter obi agemate agbay...	-1.665335e-17		Negative
454	noticed kwankwaso vanished since peter obi sup...	-2.775558e-17		Negative
2531	much u amp love support peter obi amp never wa...	-2.775558e-17		Negative
9212	dont think people fully understand means citiz...	-3.700743e-17		Negative

```
In [28]: tweet = ' '.join([word for word in neg_tweet["Tweets"]])
plt.figure(figsize=(20,14),facecolor='None')
wordcloud=WordCloud(max_words=500,width=1800,height=1000).generate(tweet)
plt.imshow(wordcloud,interpolation='bilinear')
plt.axis('off')
plt.title('Most Frequent Word in Negative Tweet',fontsize=19)
plt.savefig("Cloud.png", format="png")
plt.show()
```

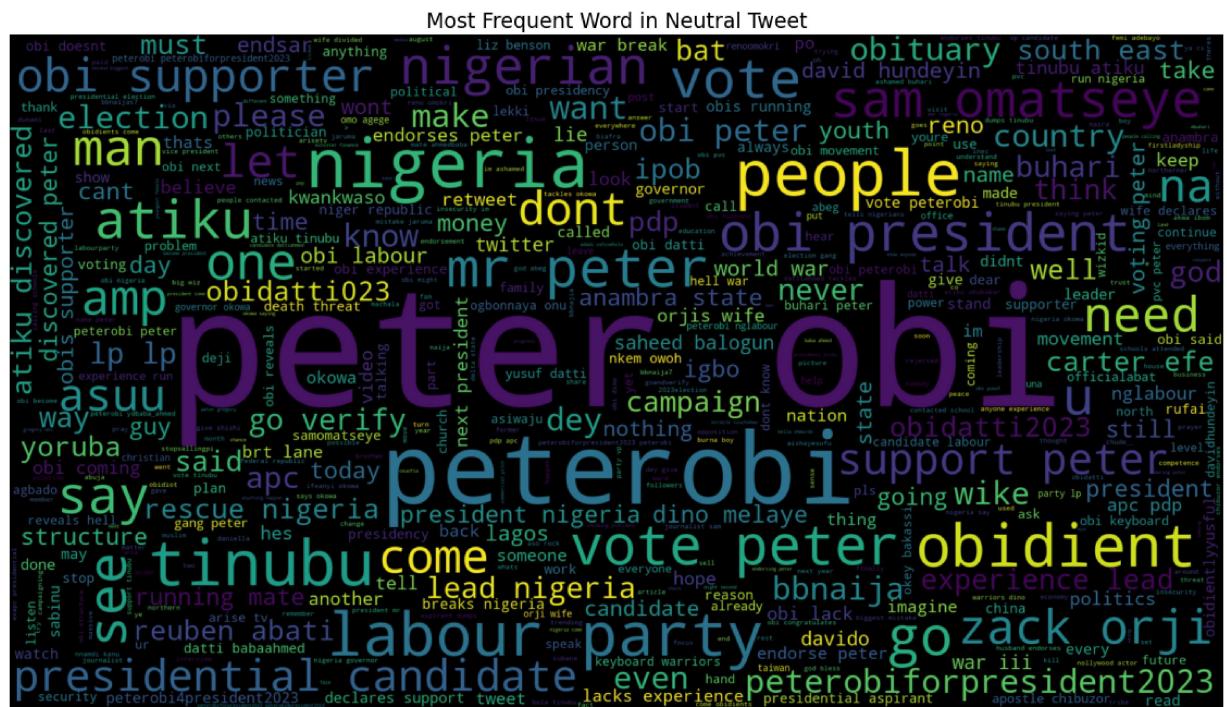


```
In [29]: neu_tweet=df[df.sentiment=="Neutral"]
neu_tweet=neu_tweet.sort_values(['polarity'], ascending=False)
neu_tweet.head()
```

Out[29]:

		Tweets	polarity	sentiment
2	im going jeffreygutermans upcoming space join ...	0.0	Neutral	
12727	happened rabiu kwankwasos presidential campaig...	0.0	Neutral	
12723	peter obi reputation religious ethnic intolerance	0.0	Neutral	
12722	youd want tear watch see people 4 peterobi ens...	0.0	Neutral	
12719	dear nigerians imagine future would looked lik...	0.0	Neutral	

```
In [40]: tweet = ' '.join([word for word in neu_tweet["Tweets"]])
plt.figure(figsize=(20,14),facecolor='None')
wordcloud=WordCloud(max_words=500,width=1800,height=1000).generate(tweet)
plt.imshow(wordcloud,interpolation='bilinear')
plt.axis('off')
plt.title('Most Frequent Word in Neutral Tweet',fontsize=19)
plt.savefig("Cloud.png3", format="png")
plt.show()
```



```
In [31]: vec= CountVectorizer(ngram_range=(1,2)).fit(df['Tweets'])
```

```
In [32]: x=df['Tweets']
          y=df['sentiment']
          x=vec.transform(x)
```

```
In [33]: x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.25,random_state=42)
```

```
In [34]: dtc=DecisionTreeClassifier(random_state=45)
```

```
In [35]: dtc.fit(x_train,y_train)
```

```
Out[35]: DecisionTreeClassifier(random_state=45)
```

```
In [36]: dtc_mypred=dtc.predict(x_test)  
print(classification_report(y_test,dtc_mypred))
```

	precision	recall	f1-score	support
Negative	0.81	0.67	0.74	772
Neutral	0.91	0.96	0.94	1961
Positive	0.89	0.90	0.90	1988
accuracy			0.89	4721
macro avg	0.87	0.85	0.86	4721
weighted avg	0.89	0.89	0.89	4721

```
In [37]: accuracy_score(y_test,dtc_mypred)*100
```

```
Out[37]: 89.09129421732683
```

```
In [ ]:
```