

## DA Lab 4

### Text Analytics of Twitter Data on Demonetization

**Name: Dheeraj Chaudhary**

**Roll: 17BCS009**

Let's consider a "Hypothesis" that majority of the population of the country are in support of Demonetization

We'll check our Hypothesis at the end using various methods

In [58]:

```
# Importing the required libraries
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from textblob import TextBlob #Library for Text Processing
from bs4 import BeautifulSoup
import re
from wordcloud import WordCloud, STOPWORDS
import networkx as nx # for visualisation
import nltk
nltk.download('vader_lexicon')
nltk.download('stopwords')
from nltk.corpus import stopwords
import itertools
import collections
from nltk import bigrams
from sklearn.feature_extraction.text import TfidfVectorizer
from nltk.stem import WordNetLemmatizer
from sklearn.cluster import KMeans
```

```
[nltk_data] Downloading package vader_lexicon to
[nltk_data] /home/dheeraj/nltk_data...
[nltk_data] Package vader_lexicon is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data] /home/dheeraj/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
```

In [8]:

```
tweets=pd.read_csv("/home/dheeraj/my_projects/my_project_env/practice/sentiment_analysis/tweets.csv")
tweets.head()
```

Out[8]:

	Unnamed: 0	x	text	favorited	favoriteCount	replyToSN	created	truncated	retweetCount
0	1	1	RT @rssurjewala: Critical question: Was PayTM ...	False	0	NaN	2016-11-23 18:40:30	False	0
1	2	2	RT @Hemant_80: Did you vote on #Demonetization...	False	0	NaN	2016-11-23 18:40:29	False	0
2	3	3	RT @roshankar: Former FinSec, RBI Dy Governor,...	False	0	NaN	2016-11-23 18:40:03	False	0
3	4	4	RT @ANI_news: Gurugram (Haryana): Post office ...	False	0	NaN	2016-11-23 18:39:59	False	0
4	5	5	RT @satishacharya: Reddy Wedding! @mail_today ...	False	0	NaN	2016-11-23 18:39:39	False	0

## Number of rows and column in the taken dataset

In [9]:

```
tweets.shape
```

Out[9]:

```
(14940, 16)
```

## Some of the tweet phrases from the dataset without cleaning

In [10]:

```
display(tweets.text.head(10))
```

```
0    RT @rssurjewala: Critical question: Was PayTM ...
1    RT @Hemant_80: Did you vote on #Demonetization...
2    RT @roshankar: Former FinSec, RBI Dy Governor,...
3    RT @ANI_news: Gurugram (Haryana): Post office ...
4    RT @satishacharya: Reddy Wedding! @mail_today ...
5    @DerekScissors1: India's #demonetization: #Bla...
6    RT @gauravcsawant: Rs 40 lakh looted from a ba...
7    RT @Joydeep_911: Calling all Nationalists to j...
8    RT @sumitbhati2002: Many opposition leaders ar...
9    National reform now destroyed even the essence...
Name: text, dtype: object
```

## Number of retweets for each of the given 14940 tweets whose indexes are given

In [11]:

```
print(tweets['retweetCount'])
```

```
0      331
1       66
2       12
3      338
4      120
...
14935    4
14936    0
14937   30
14938    2
14939    2
Name: retweetCount, Length: 14940, dtype: int64
```

**Now clean the dataset by removing HTML tags using BeautifulSoup, removing non letters present in the tweets and after doing this, split the text and remove the stowords present in the tweets and then rejoin the tweets**

In [18]:

```
def clean(x):
    #Remove Html
    x=BeautifulSoup(x).get_text()

    #Remove Non-Letters
    x=re.sub('[^a-zA-Z]', ' ',x)

    #Convert to lower_case and split
    x=x.lower().split()

    #Remove stopwords
    stop=set(stopwords.words('english'))
    words=[w for w in x if not w in stop]

    #join the words back into one string
    return(' '.join(words))
tweets['text']=tweets['text'].apply(lambda x:clean(x))
```

## Some of the tweets phrases after cleaning

In [19]:

```
# tweets.head()
display(tweets.text.head(10))
```

```
0    rt rssurjewala critical question paytm informe...
1        rt hemant vote demonetization modi survey app
2    rt roshankar former finsec rbi dy governor cbd...
3    rt ani news gurugram haryana post office emplo...
4    rt satishacharya reddy wedding mail today cart...
5    derekscissors india demonetization blackmoney ...
6    rt gauravcsawant rs lakh looted bank kishtwar ...
7    rt joydeep calling nationalists join walk corr...
8    rt sumitbhati many opposition leaders narendra...
9    national reform destroyed even essence sagan i...
Name: text, dtype: object
```

## Create a list of lists containing lowercase words for each tweet

In [15]:

```
words_in_tweet = [tweets.lower().split() for tweets in tweets.text]  
words_in_tweet
```

Out[15]:

```
[['rt',  
  'rssurjewala',  
  'critical',  
  'question',  
  'paytm',  
  'informed',  
  'demonetization',  
  'edict',  
  'pm',  
  'clearly',  
  'fishy',  
  'requires',  
  'full',  
  'disclosure'],  
 ['rt', 'hemant', 'vote', 'demonetization', 'modi', 'survey', 'ap  
p'],  
 ['rt',  
  'roshankar']]
```

**For additional clean-up, as to remove words that do not add meaningful information to the text you are trying to analysis. These words referred to as “stop words” and include commonly appearing words such as who, what, you, etc.**

In [22]:

```
stop_words = set(stopwords.words('english'))  
  
# View a few words from the set  
list(stop_words)[0:10]
```

In [23]:

```
tweets_nsw = [[word for word in tweet_words if not word in stop_words]  
               for tweet_words in words_in_tweet]
```

In [24]:

```
tweets_nsw
'at',
'dy',
'governor',
'cbdt',
'chair',
'harvard',
'professor',
'lambaste',
'demonetization',
'aam',
'aadmi',
'listen',
'th'],
['rt',
'ani',
'news',
'gurugram',
'haryana',
'post',
'office',
,
```

**Again, i'm creatig a counter to return the most commonly used words and the number of times that they are used in the tweet**

In [ ]:

```
all_words_nsw = list(itertools.chain(*tweets_nsw))
counts_nsw = collections.Counter(all_words_nsw)
counts_nsw.most_common(15)
```

**Plotting the word frequencies after removing the stop words**

In [27]:

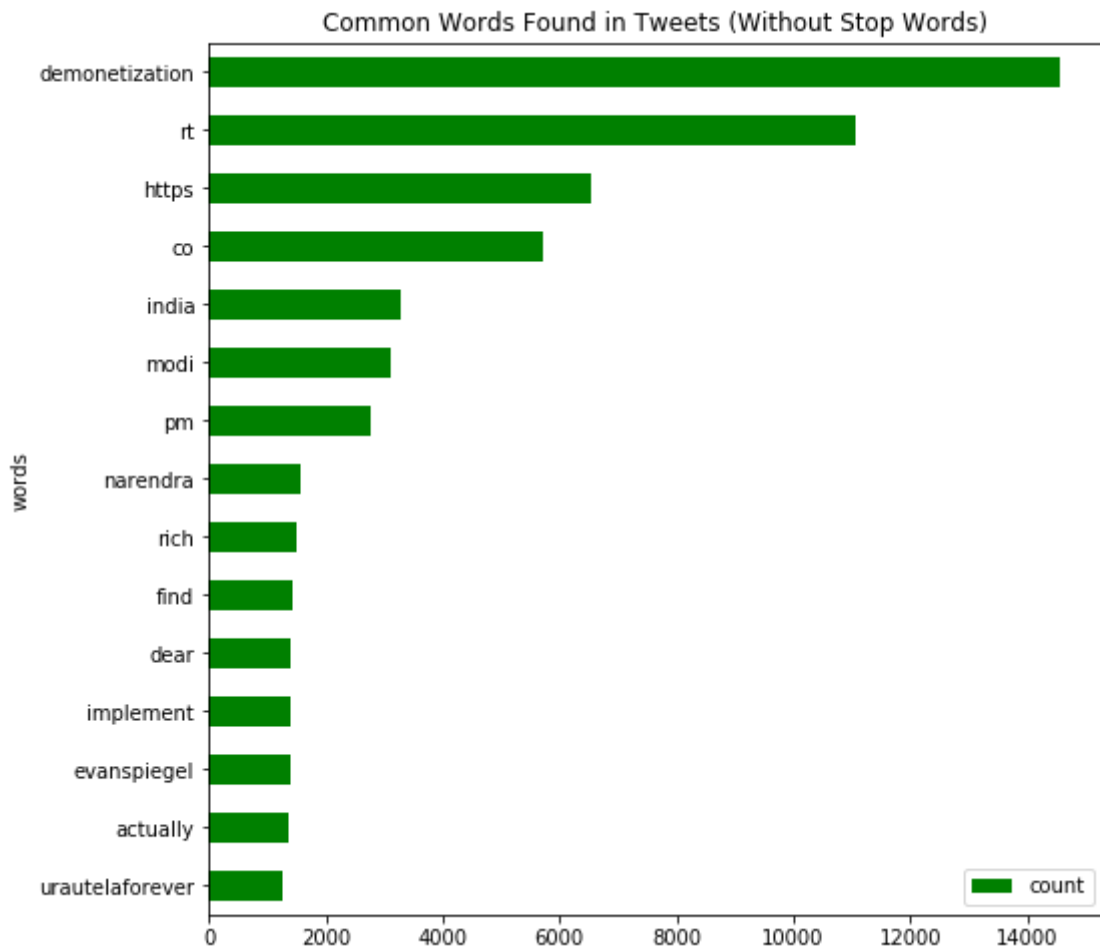
```
clean_tweets_nsw = pd.DataFrame(counts_nsw.most_common(15),
                                columns=['words', 'count'])

fig, ax = plt.subplots(figsize=(8, 8))

# Plot horizontal bar graph
clean_tweets_nsw.sort_values(by='count').plot.barh(x='words',
                                                    y='count',
                                                    ax=ax,
                                                    color="green")

ax.set_title("Common Words Found in Tweets (Without Stop Words)")

plt.show()
```



## Remove some other words which are occurring in almost every tweets

In [30]:

```
collection_words = ['rt', 'co', 'http', 'https', 'j', 'k']
tweets_nsw_nc = [[w for w in word if not w in collection_words]
                  for word in tweets_nsw]
tweets_nsw_nc
```

'euct',  
 'pm',  
 'clearly',  
 'fishy',  
 'requires',  
 'full',  
 'disclosure'],  
 ['hemant', 'vote', 'demonetization', 'modi', 'survey', 'app'],  
 ['roshankar',  
 'former',  
 'finsec',  
 'rbi',  
 'dy',  
 'governor',  
 'cbdt',  
 'chair',  
 'harvard',  
 'professor',  
 'lambaste',  
 'demonetization',

## Exploring some co-occurring words using bi-grams



In [32]:

```
terms_bigram = [list(bigrams(tweet)) for tweet in tweets_nsw_nc]
terms_bigram
[('fishy', 'requires'),
 ('requires', 'full'),
 ('full', 'disclosure')],
[('hemant', 'vote'),
 ('vote', 'demonetization'),
 ('demonetization', 'modi'),
 ('modi', 'survey'),
 ('survey', 'app')],
[('roshankar', 'former'),
 ('former', 'finsec'),
 ('finsec', 'rbi'),
 ('rbi', 'dy'),
 ('dy', 'governor'),
 ('governor', 'cbdt'),
 ('cbdt', 'chair'),
 ('chair', 'harvard'),
 ('harvard', 'professor'),
 ('professor', 'lambaste'),
 ('lambaste', 'demonetization'),
 ('demonetization', 'aam'),
```

## Counting top 20 common bi-grams

In [35]:

```
# Flatten list of bigrams in clean tweets
bigrams = list(itertools.chain(*terms_bigram))

# Create counter of words in clean bigrams
bigram_counts = collections.Counter(bigrams)

bigram_counts.most_common(20)
```

Out[35]:

```
[(('narendra', 'modi'), 1566),
 (('india', 'rich'), 1431),
 (('pm', 'narendra'), 1424),
 (('demonetization', 'find'), 1399),
 (('implement', 'demonetization'), 1398),
 (('rich', 'pm'), 1391),
 (('modi', 'implement'), 1376),
 (('evanspiegel', 'india'), 1371),
 (('dear', 'evanspiegel'), 1366),
 (('find', 'actually'), 1311),
 (('urautelaforever', 'dear'), 1273),
 (('narendramodi', 'demonetization'), 595),
 (('pm', 'narendramodi'), 590),
 (('since', 'demonetization'), 572),
 (('rs', 'lakh'), 551),
 (('demonetization', 'terrorists'), 550),
 (('lakh', 'looted'), 544),
 (('looted', 'bank'), 544),
 (('bank', 'kishtwar'), 542),
 (('kishtwar', 'third'), 542)]
```

## creating a dataframe for visualization of bigrams

In [36]:

```
#Top 20 most common bigrams
bigram_df = pd.DataFrame(bigram_counts.most_common(20),
                          columns=['bigram', 'count'])
bigram_df
```

Out[36]:

	bigram	count
0	(narendra, modi)	1566
1	(india, rich)	1431
2	(pm, narendra)	1424
3	(demonetization, find)	1399
4	(implement, demonetization)	1398
5	(rich, pm)	1391
6	(modi, implement)	1376
7	(evanspiegel, india)	1371
8	(dear, evanspiegel)	1366
9	(find, actually)	1311
10	(urauteleforever, dear)	1273
11	(narendramodi, demonetization)	595
12	(pm, narendramodi)	590
13	(since, demonetization)	572
14	(rs, lakh)	551
15	(demonetization, terrorists)	550
16	(lakh, looted)	544
17	(looted, bank)	544
18	(bank, kishtwar)	542
19	(kishtwar, third)	542

## Visualizing network of bi-grams using networkx package of python

In [39]:

```
d = bigram_df.set_index('bigram').T.to_dict('records')
G = nx.Graph()
# Create connections between nodes
for k, v in d[0].items():
    G.add_edge(k[0], k[1], weight=(v * 3))

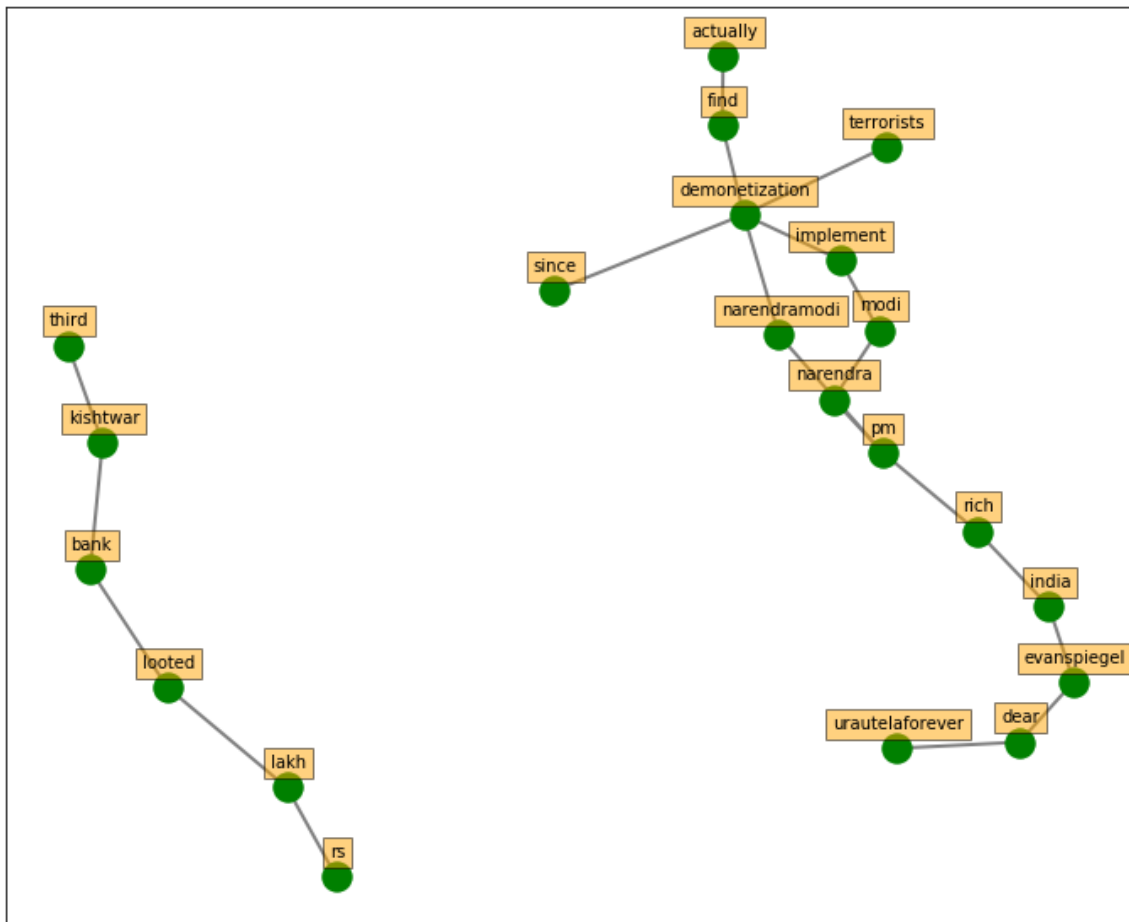
fig, ax = plt.subplots(figsize=(12, 10))

pos = nx.spring_layout(G, k=4)

# Plot networks
nx.draw_networkx(G, pos,
                  font_size=11,
                  fontweight='bold',
                  width=2,
                  edge_color='grey',
                  node_color='green',
#                  edge_length = 10,
                  with_labels = False,
                  ax=ax)

# Create offset labels
for key, value in pos.items():
    x, y = value[0]+.00167, value[1]+.045
    ax.text(x, y,
            s=key,
            bbox=dict(facecolor='orange', alpha=0.5),
            horizontalalignment='center', fontsize=10)

plt.show()
```



In [ ]:

```
#TextBlob("not great").sentiment
```

## Sentiment Analysis using nltk

In [40]:

```

from nltk.sentiment.vader import SentimentIntensityAnalyzer
from nltk.sentiment.util import *
from nltk.corpus import stopwords
from nltk import tokenize

sid = SentimentIntensityAnalyzer()

tweets['sentiment_compound_polarity']=tweets.text.apply(lambda x:sid.polarity_score
tweets['sentiment_neutral']=tweets.text.apply(lambda x:sid.polarity_scores(x)['neu']
tweets['sentiment_negative']=tweets.text.apply(lambda x:sid.polarity_scores(x)['neg']
tweets['sentiment_pos']=tweets.text.apply(lambda x:sid.polarity_scores(x)['pos'])
tweets['sentiment_type']=''
tweets.loc[tweets.sentiment_compound_polarity>0,'sentiment_type']='Positive'
tweets.loc[tweets.sentiment_compound_polarity==0,'sentiment_type']='Neutral'
tweets.loc[tweets.sentiment_compound_polarity<0,'sentiment_type']='Negative'
tweets.head()

```

Out[40]:

Unnamed: 0	x	text	favorited	favoriteCount	replyToSN	created	truncated	reply1
0	1	rt rssurjewala critical question paytm informe...	False	0	NaN	2016-11-23 18:40:30	False	
1	2	rt hemant vote demonetization modi survey app	False	0	NaN	2016-11-23 18:40:29	False	
2	3	rt roshankar former finsec rbi dy governor cbd...	False	0	NaN	2016-11-23 18:40:03	False	
3	4	rt ani news gurugram haryana post office emplo...	False	0	NaN	2016-11-23 18:39:59	False	
4	5	rt satishacharya reddy wedding mail today cart...	False	0	NaN	2016-11-23 18:39:39	False	

5 rows × 21 columns

## Counting total number of tweets for each category of sentiment

**We can see number of tweets having Positive sentiment is 6771, neutral sentiment is 4233 and negative sentiment is 3936**

In [25]:

```
tweets.sentiment_type.value_counts()
```

Out[25]:

```
Positive    6771  
Neutral     4233  
Negative    3936  
Name: sentiment_type, dtype: int64
```

## Bar Plot for Sentiment Analysis

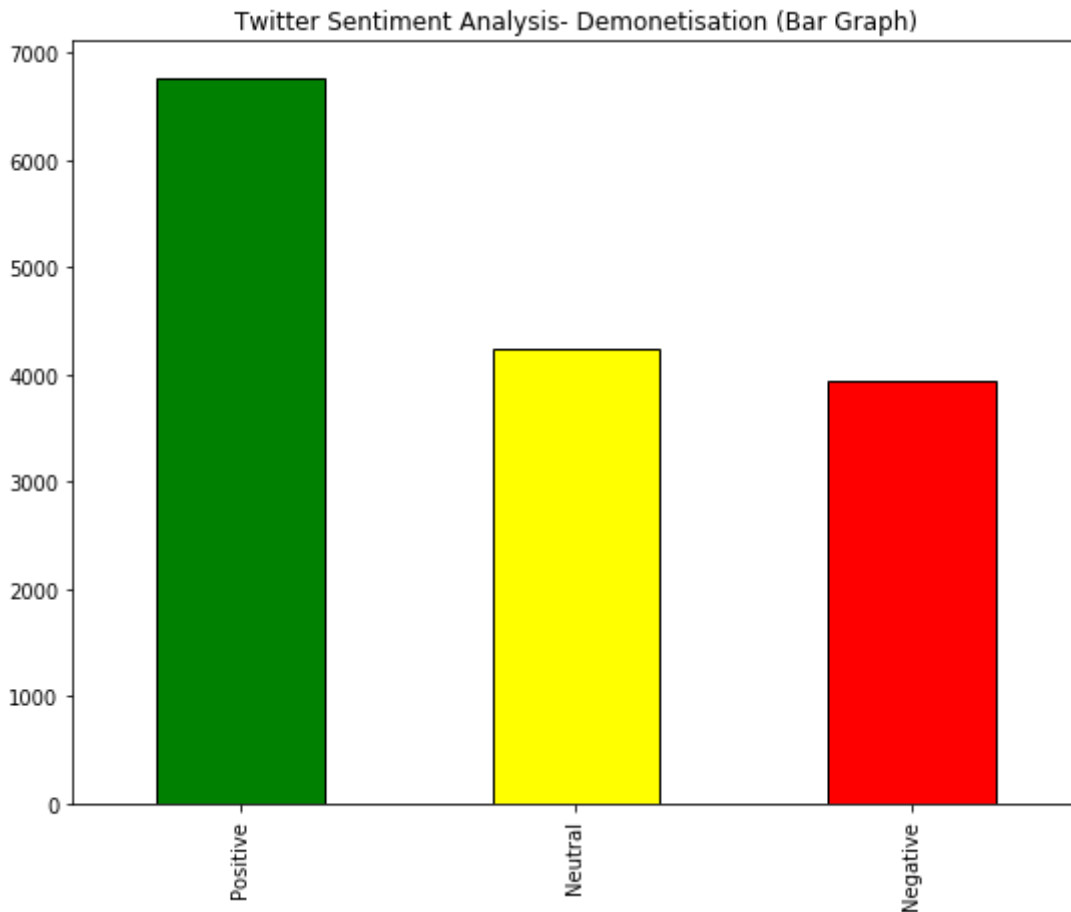
**We can see Positive sentiment is higher than negative sentiment**

In [26]:

```
# fig = plt.figure()
# fig.savefig('Sentiment_bar_plot.pdf')
colors = ['green', 'yellow', 'red']
tweets.sentiment_type.value_counts().plot(kind='bar',figsize=(9, 7),edgecolor='k',t
# plt.savefig('Sentiment_bar_plot.png', dpi=100)
```

Out[26]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7f2894a0be80>



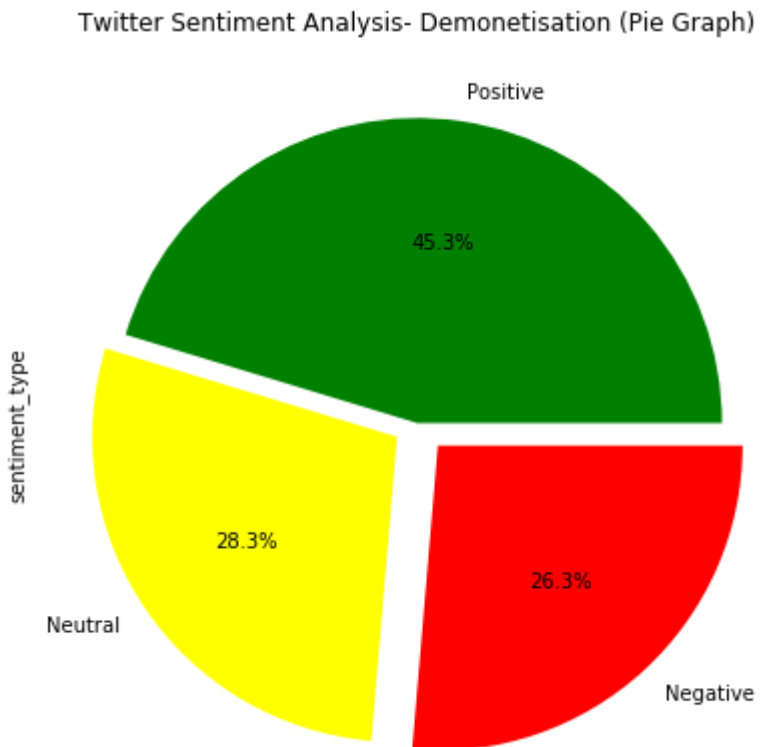
## Sentiment plot in forms of Pie chart

In [27]:

```
colors = ['green', 'yellow', 'red']  
explode = (0, 0.08, 0.1)  
tweets.sentiment_type.value_counts().plot(kind='pie', figsize=(9, 7), title="Twitter  
# plt.savefig('Sentiment_pie_plot.png', dpi=100)
```

Out[27]:

<matplotlib.axes.\_subplots.AxesSubplot at 0x7f2894ac5198>



## Source of tweets

**We can see the source of tweets corresponding to each index of all 14940 tweets**

**From the below pie chart it's clearly visible that, Top 3 of Source: 1 - Twitter For Android 2 - Twitter for Iphones 3 - Twitter for Web Client and finally !**



In [47]:

```

tweets['statusSource_new'] = ''

for i in range(len(tweets['statusSource'])):
    m = re.search('(?!<=>)(.*)', tweets['statusSource'][i])
    try:
        tweets['statusSource_new'][i]=m.group(0)
    except AttributeError:
        tweets['statusSource_new'][i]=tweets['statusSource'][i]

#print(tweets['statusSource_new'].head())

tweets['statusSource_new'] = tweets['statusSource_new'].str.replace('</a>', ' ', ca
tweets['statusSource_new2'] = ''

for i in range(len(tweets['statusSource_new'])):
    if tweets['statusSource_new'][i] not in ['Twitter for Android ', 'Twitter Web Cl
        tweets['statusSource_new2'][i] = 'Others'
    else:
        tweets['statusSource_new2'][i] = tweets['statusSource_new'][i]
print(tweets['statusSource_new2'])

tweets_by_type2 = tweets.groupby(['statusSource_new2'])['retweetCount'].sum()
tweets_by_type2.rename("", inplace=True)
explode = (0, 0, 0.09, 0)
tweets_by_type2.transpose().plot(kind='pie', figsize=(9, 7), autopct='%1.1f%%', shadow
plt.legend(bbox_to_anchor=(1, 1), loc=6, borderaxespad=0.)
plt.title('Number of retweetcount by Source bis', bbox={'facecolor':'0.8', 'pad':5})

```

/home/dheeraj/my\_projects/my\_project\_env/lib/python3.6/site-packages/i  
pykernel\_launcher.py:6: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: [http://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) (h  
[http://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#re  
turning-a-view-versus-a-copy](http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

/home/dheeraj/my\_projects/my\_project\_env/lib/python3.6/site-packages/i  
pykernel\_launcher.py:19: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: [http://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) (h  
[http://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#re  
turning-a-view-versus-a-copy](http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

/home/dheeraj/my\_projects/my\_project\_env/lib/python3.6/site-packages/i  
pykernel\_launcher.py:17: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: [http://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) (h  
[http://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#re  
turning-a-view-versus-a-copy](http://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

```

0      Twitter for Android
1      Twitter for Android
2      Twitter for Android
3      Twitter for Android

```

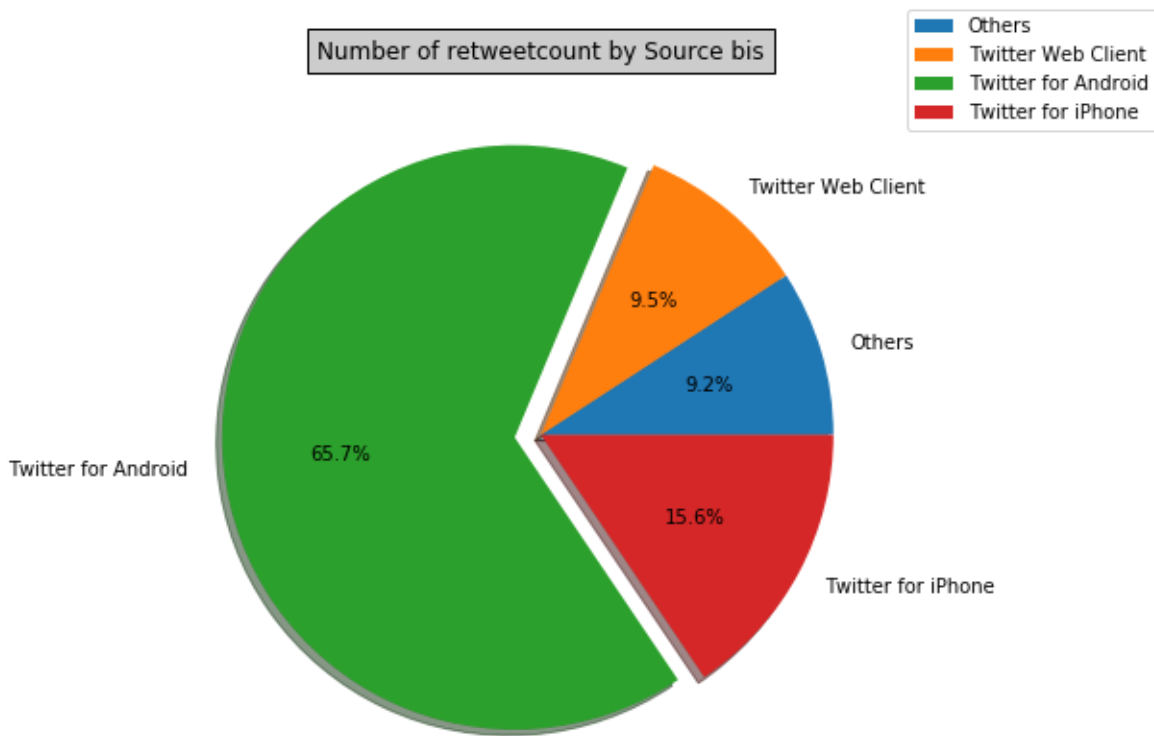
```
4                                Others
```

```
...
14935  Twitter for Android
14936  Twitter for Android
14937  Twitter for Android
14938  Twitter Web Client
14939  Twitter Web Client
```

```
Name: statusSource_new2, Length: 14940, dtype: object
```

```
Out[47]:
```

```
Text(0.5, 1.0, 'Number of retweetcount by Source bis')
```



## Clusterig using K-means algorithm

**We can see the shape of X and 100 clusters generated usig above K-means algorithm**

In [66]:

```

tweets['text_lem'] = [''.join([WordNetLemmatizer().lemmatize(re.sub('[^A-Za-z]', '
####
vectorizer = TfidfVectorizer(max_df=0.5,max_features=10000,min_df=10,stop_words='en
X = vectorizer.fit_transform(tweets['text_lem'].str.upper())
print("Shape of X is: ", X.shape, '\n')
print("Clusters are: ", '\n')
km = KMeans(n_clusters=100,init='k-means++',max_iter=200,n_init=1)
km.fit(X)
terms = vectorizer.get_feature_names()
order_centroids = km.cluster_centers_.argsort()[:,::-1]
for i in range(100):
    print("cluster %d:" %i, end='')
    for ind in order_centroids[i,:10]:
        print(' %s' % terms[ind], end='')
    print()

```

Shape of X is: (14940, 1925)

Clusters are:

cluster 0: kishtwar gauravcsawant incident looted terrorists lakh rs b  
ank https gau

cluster 1: huge su voter modibharosa nation support narendramodi peopl  
e pm https

cluster 2: oscar obqrhlsl drkumarvishwas goes mr https kumar retweete  
d dr hazare

cluster 3: effect https said surrender naxals weekend supplying requir  
ed sources aapnewsalert

cluster 4: prove successful wave adityanath yogi polls result today ra  
hulroushan modi

cluster 5: poo aanupriyaaa actually evanspiegel implement dear rich n  
arendra pm modi

cluster 6: retaining byelection proved seats abov ls ippatel state bjp  
support

cluster 7: putting nitishkumar uodwxdpmmg party politics supports modi  
bharosa nation narendramodi pm

cluster 8: glvmi munaf ages patel seeing roflgandhi thought lost funct  
ion fund

cluster 9: flags watched reinstate ge timcast ads videos youtube peopl  
e money

cluster 10: hemant vote survey app modi positive got https pm future

cluster 11: goalpost royal toss indiaexplained cag audit paresh sanjay  
uv hi going

cluster 12: englis gutkha tobacco prohibition banned quiet love men ke  
pt rahulroushan

cluster 13: wants views narendramodi pm https app bypolls banknotes pr  
inted narendra

cluster 14: edict informed critical rssurjewala fishy requires disclos  
ure clearly paytm question

cluster 15: urautelaforever implement actually dear evanspiegel rich n  
arendra pm modi india

cluster 16: ujjwala urea coated neem mandis historically dhan unpreced  
ent jan impact

cluster 17: censorship adsense youtube boycott truth https alert googl  
e red video

cluster 18: coast fishermen kerala parliament shashitharoor question i  
mpact india https retweeted

cluster 19: dire reconfigured bulletin continues severe liquidity barb  
arindian bcoz useless crisis

cluster 20: loosened restrictions withdrawal hanke steve demo claims r  
unning atms cash

cluster 21: education overall motilal oswal services article impact ht  
tps india ag

cluster 22: urjitpatelrbi poll demonetisation rbi think notes modi ret  
weeted patel news

cluster 23: duty anandkamal frankly wil parlmnt address understand sha  
shitharoor narendramodi disruption

cluster 24: effects https incrementpremierleague miller john glr marke  
t common ban note

cluster 25: milkman tarekfatah adityanair rashminp prasadaol asked che  
ck years time narendramodi

cluster 26: corruptionfreeindia walk nationalists spread joydeep mas c  
alling benefits join https

cluster 27: ends daughter gujarat fund wedding life ahead man shortage  
dgai

cluster 28: money black https reason white lot given unaffected people  
folks

cluster 29: pmoindia minimalism taught happiness thank stuff drshobha  
insights prime minister

cluster 30: headline closing lack staff attomeybharti blames supply ye  
ar report end

cluster 31: bank asian acc priyaakulkarni positiv development grow wor  
ld ban note

cluster 32: secy dea briefing morning dasshaktikanta tdj pib watch uo  
india

cluster 33: nhek tpg heard gandhi atheist krishna standing ji https gg  
j

cluster 34: finds subtitles hitler gkiugm rng hilarious themehulpatel  
read video watch

cluster 35: whooping profit ak windfall pappu crores kanimozhi mamata  
days like

cluster 36: mallya bailed pet vijaymallya loan waived got modi dog wit  
hinc

cluster 37: index whammy consumer industrial double production iip con  
tracts cpi mahcongress

cluster 38: bibi prob minimathur explains tweet drgpradhan says https  
gandhi gau

cluster 39: opinions lacs harshkkapoor amazing hours commitment na fig  
ht results response

cluster 40: youtube https video liked turks young future google patreo  
n scot

cluster 41: suspect overwhelming elec political far positive economic  
going impact gauravcsawant

cluster 42: scale large finally gone citizens country ahead centerofri  
ght pain funds

cluster 43: demonstrating solidarity joined demanding mps parliame lin  
e address opposition pmoindia

cluster 44: wgdeaa begun revert rlr morgan jp data milanv mean economi  
c

cluster 45: heritage shares sold prior evident ncbn announcement knowl  
edge ysraparty days

cluster 46: swift deployment lead worldbank technology secretarymeity  
payments http positive digital

cluster 47: good dying agenda op run ppl banks black money gau

cluster 48: paan owner shop gets ahmedabad fake note rs epvcu dna

cluster 49: feku forgetting motive pro exercise means miteshpattel pr d  
rama media

cluster 50: vindicated parliamenta historic honest feel indiafightscor

ruption thore ra arunjaitley ji  
cluster 51: seizes force malda border security worth bengal fake under stand askanshul  
cluster 52: notes new old rs https getting calibrated currency dispens e centerofright  
cluster 53: https india narendramodi modi cash people like pm indian n ews  
cluster 54: opks chamchas wiped terrorism told aartic modi https gadg e ts future  
cluster 55: bummed frrvmo acq conversation missed cgdev annielowrey mi lanv arvindsubraman video  
cluster 56: vaidyanathan prof mudra action writes come hkkdva upma uns time  
cluster 57: pygk rmg joydas strategy feedback taking app question supp ort people  
cluster 58: heteyum announcements mib finminindia ch dasshaktikanta in dia https gandhi gau  
cluster 59: tp svf urautelaforever evanspiegel implement dear rich nar endra pm modi  
cluster 60: momentous courage immense needed indiafightscorruption tho re ra arunjaitley decision narendramodi  
cluster 61: tjain execut planned corruption decision good narendramodi zt gates gauravcsawant  
cluster 62: realised plastic moving families hand trxn s live online ce nterofright cash  
cluster 63: misery mounting indifferent totally proves hards cpiml dip ankar app govt  
cluster 64: oppn netascashin uv ucs away wants timesnow debate running govt  
cluster 65: creativity umbvq fmh productivity realhistorypic increased say https funds future  
cluster 66: batch watching nomoneyyaar maoists comes dipendradipzo com e effect https gave  
cluster 67: slip appearing offensive youtubeisdead percent markdice th ads videos youtube  
cluster 68: payments digital catalyzing growth https india mobile fint ech mo se  
cluster 69: delhities yesiamsaffron cursing thug que praising htt long standing arvindkejriwal  
cluster 70: anurag finances squeezed raided tight old new jan notes te rrorists  
cluster 71: cr lakhs mahikainfra returned govt vote days bank ba phase  
cluster 72: good reflected transparent actions piyushgoyaloffc makes e conomics believe politics future  
cluster 73: lifeinsurance wise hour secure bajajallianzlic initiative need think ahead sc  
cluster 74: forecast imf annual growth points cnnnews impact india htt ps ahead  
cluster 75: gau rakshaks drives swachata stunts diwas yoga romeo mkatj u voted  
cluster 76: bank robberies routine army loot kashmir daily askanshul t errorists security  
cluster 77: https india modi people common impact actually demonetisat ion economy indian  
cluster 78: nsitharaman narendramodi https zt gave fund funding funds future gadgets  
cluster 79: recalibrated amoun small majority coming ppl new notes cen terofright atms  
cluster 80: continuous struggle value beginning high deep currency gov ernment black pragra  
cluster 81: nefarious designs nexus kashmirunrest dealt separatist blo

w jammukashmir gui body  
cluster 82: zh atheist krishna effect https zt gauravcsawant gau gates  
gandhi  
cluster 83: recovered reports miteshpatel pr drama blackmoney https ga  
tes gauravcsawant gau  
cluster 84: pocketing pick eqbmve guise charges service jamewils https  
gates gandhi  
cluster 85: boycottsnapchat shiralkar nr bc poor actually rich modi in  
dia implement  
cluster 86: ratantata frthy ltth supports modibharosa narendra pm modi  
https funding  
cluster 87: telangana innovative solution minit slccpr nvv farmers hel  
p affected govt  
cluster 88: internetfreedom collector jackerhack indore shut problem t  
hink like https future  
cluster 89: debts amounting queued gauravpandhi outside bad waived cro  
res mallya banks  
cluster 90: rebel politik succinctly bspup invi mayawatinextupcm expla  
ined notenahipmbadlo bsp pm  
cluster 91: unsubtledesit posts pic dead death kejriwal related claims  
man https  
cluster 92: exam teen fees withdraw hangs unable iiielppllf mmgps bsind  
ia uo  
cluster 93: showrooms firing quietly smaller executives sales shops sa  
sumitkashyapjha started  
cluster 94: lord behen bagwati certainly hurting mamtaofficialai officeo  
frg shefvaitya lot arvindkejriwal  
cluster 95: ybf taeyns failure rahulroushan https frustindian function  
fund funding funds  
cluster 96: agricultural panvel board body matter rupasubramanya big p  
olls lost time  
cluster 97: compromised manner vocal ag seriously demon finally pan me  
meghnad government  
cluster 98: bankers hopefully return normalcy ava soon dry months desp  
ite post  
cluster 99: panic mode creating running unrest civil cash india https  
youtube

### Printing out 100 clusters

In [62]:

```
cluster 0: kishwar gauravcsawant incident looted terrorists lakh rs
bank https gau
cluster 1: tp svf urautelaforever evanspiegel implement dear rich na
rendra pm modi
cluster 2: pygk rmj joydas strategy feedback taking app question sup
port people
cluster 3: huge su voter modibharosa nation support narendramodi peo
ple pm https
cluster 4: oscar obqrhlsl drkumarvishwas goes mr https kumar retwee
ted dr hazare
cluster 5: ujjwala urea coated neem mandis historically dhan unprece
dent jan impact
cluster 6: urautelaforever implement actually dear evanspiegel rich
narendra pm modi india
cluster 7: suspect overwhelming elec political far positive economic
going impact gauravcsawant
cluster 8: https india modi people common effect impact actually dem
onetisation economy
cluster 9: loosened restrictions withdrawal hanke steve demo claims
running atms cash
cluster 10: retaining byelection proved seats abov ls ippatel state
bjp support
cluster 11: whooping profit ak windfall pappu crores kanimozhi mamat
a days like
cluster 12: blackmoney recovered reports miteshpattel pr drama https
india narendramodi dept
cluster 13: glvmi munaf ages patel seeing roflgandhi thought lost fu
nction fund
cluster 14: edict informed critical rssurjewala fishy requires discl
osure clearly paytm question
cluster 15: money black https reason white lot given unaffected peop
le folks
cluster 16: good reflected transparent actions piyushgoyaloffc makes
economics believe politics future
cluster 17: parliament detaiend station msisodia street police cm ma
rch dy daarubaazmehta
cluster 18: nhek tpg heard gandhi atheist krishna standing ji https
ggj
cluster 19: gau rakshaks drives swachata stunts diwas yoga romeo mka
tju voted
cluster 20: agricultural panvel board body matter rupasubramanya big
polls lost time
cluster 21: minimalism happiness taught thank stuff drshobha insight
s prime minister pmoindia
cluster 22: urjitpatelrbi poll demonetisation rbi think notes modi r
etweeted patel news
cluster 23: times really helped lot growth going digital pa https ga
ndhi
cluster 24: heteyum announcements mib finminindia ch dasshaktikanta
india https gandhi gau
cluster 25: blog started years youtube people zt gau gates gandhi ga
dgets
cluster 26: putting nitishkumar uodwxdpmmg party politics supports m
odibharosa nation narendramodi pm
cluster 27: panic mode creating running unrest civil cash india http
s youtube
cluster 28: demonstrating solidarity joined demanding mps parliame l
```

ine address opposition pmoindia  
cluster 29: flags watched reinstate ge timcast ads videos youtube pe  
ople money  
cluster 30: index whammy consumer industrial double production iip c  
ontracts cpi mahcongress  
cluster 31: ends daughter gujarat fund wedding life ahead man shorta  
ge dgai  
cluster 32: dire reconfigured bulletin continues severe liquidity ba  
rbarindian bcoz useless crisis  
cluster 33: bypolls decide madskak demonetizat results modi https fu  
nding funds fund  
cluster 34: hrs punitspeaks meradeshbadalraha response far ban note  
result survey pmoindia  
cluster 35: education overall motilal oswal services article impact  
https india ag  
cluster 36: https india youtube narendramodi cash modi people govt p  
m news  
cluster 37: loopholes sagarikaghose bdutt ranaayyub amiteshk pak pow  
er getting years bjp  
cluster 38: bank asian acc priyaakulkarni positiv development grow w  
orld ban note  
cluster 39: united moment unsubtledes i hit opposition called bharat  
nov protesting creating  
cluster 40: food vouchers attempt turning cashless digital india htt  
ps nk ng  
cluster 41: bank robberies routine army loot kashmir daily askanshul  
terrorists security  
cluster 42: paan owner shop gets ahmedabad fake note rs epvcu dna  
cluster 43: ts bandwagon ti opposing stop exactly pappu kanimozhi op  
position modi  
cluster 44: finds subtitles hitler gkiugm rng hilarious themehulpate  
l read video watch  
cluster 45: indianoilcl hinterland mini indiakadil ioc helping peopl  
e giving karnataka rural  
cluster 46: prove successful wave adityanath yogi polls result today  
rahulroushan modi  
cluster 47: extn hoarded accts depositing expect guess start hv ppl  
centerofright  
cluster 48: coast fishermen kerala parliament shashitharoor question  
impact india https retweeted  
cluster 49: bidmvfm humorministry sonunigam https zt gave funding fu  
nds future gadgets  
cluster 50: creativity umbvq fmh productivity realhistorypic increas  
ed say https funds future  
cluster 51: opks chamchas wiped terrorism told aartic modi https gad  
gets future  
cluster 52: new hnis convert cut bm calls means getting heard notes  
cluster 53: zh atheist krishna effect https zt gauravcsawant gau gat  
es gandhi  
cluster 54: delhitie yesiamsaffron cursing thug que praising htt lo  
ng standing arvindkejriwal  
cluster 55: englis gutkha tobacco prohibition banned quiet love men  
kept rahulroushan  
cluster 56: nvey secy dea dasshaktikanta briefing morning watch finm  
inindia https pib  
cluster 57: mallya bailed pet vijaymallya loan waived got modi dog w  
ithinc  
cluster 58: itd widening deepening base opportunities created income  
taxindia tax used cr  
cluster 59: swift deployment lead worldbank technology secretarymeit  
y payments http positive digital



cluster 60: bibi prob minimathur explains tweet drgpradhan says http  
s gandhi gau  
cluster 61: yarmolukdan tyriiekziu riffin https zt frustindian funct  
ion fund funding funds  
cluster 62: heritage shares sold prior evident ncbn announcement kno  
wledge ysrcparty days  
cluster 63: rmd fpfaw convenience minor friends shirishkunder vijay  
major inconvenience mallya  
cluster 64: milkman tarekfatah adityanair rashminp prasadaol asked c  
heck years time narendramodi  
cluster 65: max suffered reaction loss anilkohli going party gates g  
auravcsawant gau  
cluster 66: tjain execut planned corruption decision good narendramo  
di zt gates gauravcsawant  
cluster 67: ysrcp tdp mohan ys ysjagan jagan friend reddy chief ncbn  
cluster 68: duty anandkamal frankly wil parlmnt address understand s  
hashitharoor narendramodi disruption  
cluster 69: headline closing lack staff attomeybharti blames supply  
year report end  
cluster 70: argument buried duplicate reminds convenientl ficn pan m  
emeghnad exactly jamewils  
cluster 71: shiva sena newindianxpress anandm sad fall line governme  
nt say decision  
cluster 72: momentous courage immense needed indiafightscorruption t  
hore ra arunjaitley decision narendramodi  
cluster 73: corruptionfreeindia walk nationalists spread joydeep mas  
calling benefits join https  
cluster 74: vindicated parliamenta historic honest feel indiafightsc  
orruption thore ra arunjaitley ji  
cluster 75: poo aanupriyaaa actually evanspiegel implement dear ric  
h narendra pm modi  
cluster 76: lifeinsurance wise hour secure bajajallianzlic initiativ  
e need think ahead sc  
cluster 77: hemant vote survey app modi positive got https pm future  
cluster 78: bummed frrvmo acq conversation missed cgdev annielowrey  
milanv arvindsubraman video  
cluster 79: slip appearing offensive youtubeisdead percent markdice  
th ads videos youtube  
cluster 80: misery mounting indifferent totally proves hards cpiml d  
ipankar app govt  
cluster 81: internetfreedom collector jackerhack indore shut problem  
think like https future  
cluster 82: seizes force malda border security worth bengal fake und  
erstand askanshul  
cluster 83: ratantata frthy ltth supports modibharosa narendra pm mo  
di https funding  
cluster 84: payments digital catalyzing growth https india mobile fi  
ntech mo se  
cluster 85: lord behen bagwati certainly hurting mamtaofficiali offic  
eofrg shefvaitya lot arvindkejriwal  
cluster 86: telangana innovative solution minit slccpr nvv farmers h  
elp affected govt  
cluster 87: pawar opposes west op turned thzav cpi kerala bengal ban  
ks  
cluster 88: lelq insist jaiposwal decision app support people narend  
ramodi modi https  
cluster 89: briefing tdj morning secy dea pib uo dasshaktikanta watc  
h india  
cluster 90: crossed fingers looking amazingatheist coming crisis end  
like gauravpandhi gauravcsawant  
cluster 91: harbhajan shagun suggestion zeenewssports struggling sin

gh gives hilarious season wedding  
cluster 92: medical admissions rot clearing colleges rupees massive  
crores saaxenanurag cash  
cluster 93: effort surveyed whopping caused inconvenience worth shef  
vaidya say people https  
cluster 94: ranjan pratim jute bose workers tea set writes thedarjch  
ron banking  
cluster 95: salaam rants sitaramyechury answered continue abhic lol  
bengal people https  
cluster 96: scale large finally gone citizens country ahead centerof  
right pain funds  
cluster 97: mass surrender naxals earlier ht dipendradipzo nomoneyya  
ar maoists coming said  
cluster 98: agu sim download jio hitler aartic ch know arvindkejriwa  
l got  
cluster 99: goalpost royal toss indiaexplained cag audit paresh sanj  
ayuv hi going

**From the above results of sentiment analysis, since number of tweets that have positive sentiment is greater than the negative and neutral sentiment.**

**So we can say that majority of population are in favour of Demonetization moreover they are supporting the government**

In [ ]: