

```
!pip install spacy
!python -m spacy download en_core_web_sm
!pip install beautifulsoup4
!pip install textblob
```

```
Requirement already satisfied: spacy in c:\users\aravind\anaconda3\
lib\site-packages (3.7.4)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in c:\
users\aravind\anaconda3\lib\site-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in c:\
users\aravind\anaconda3\lib\site-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (1.0.10)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (2.0.8)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (3.0.9)
Requirement already satisfied: thinc<8.3.0,>=8.2.2 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (8.2.3)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (1.1.2)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (2.4.8)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.4.0,>=0.1.0 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (0.3.4)
Requirement already satisfied: typer<0.10.0,>=0.3.0 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (0.9.4)
Requirement already satisfied: smart-open<7.0.0,>=5.2.1 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (5.2.1)
Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (4.65.0)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (2.31.0)
Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in
c:\users\aravind\anaconda3\lib\site-packages (from spacy) (1.10.8)
Requirement already satisfied: jinja2 in c:\users\aravind\anaconda3\
lib\site-packages (from spacy) (3.1.2)
Requirement already satisfied: setuptools in c:\users\aravind\
anaconda3\lib\site-packages (from spacy) (68.0.0)
Requirement already satisfied: packaging>=20.0 in c:\users\aravind\
anaconda3\lib\site-packages (from spacy) (23.1)
Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in c:\users\
aravind\anaconda3\lib\site-packages (from spacy) (3.4.0)
Requirement already satisfied: numpy>=1.19.0 in c:\users\aravind\
anaconda3\lib\site-packages (from spacy) (1.24.3)
Requirement already satisfied: language-data>=1.2 in c:\users\aravind\
anaconda3\lib\site-packages (from langcodes<4.0.0,>=3.2.0->spacy)
(1.2.0)
```

Requirement already satisfied: typing-extensions>=4.2.0 in c:\users\aravind\anaconda3\lib\site-packages (from pydantic!=1.8,! =1.8.1,<3.0.0,>=1.7.4->spacy) (4.7.1)

Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\aravind\anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0->spacy) (2.0.4)

Requirement already satisfied: idna<4,>=2.5 in c:\users\aravind\anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0->spacy) (3.4)

Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\aravind\anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0->spacy) (1.26.16)

Requirement already satisfied: certifi>=2017.4.17 in c:\users\aravind\anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0->spacy) (2023.7.22)

Requirement already satisfied: blis<0.8.0,>=0.7.8 in c:\users\aravind\anaconda3\lib\site-packages (from thinc<8.3.0,>=8.2.2->spacy) (0.7.11)

Requirement already satisfied: confection<1.0.0,>=0.0.1 in c:\users\aravind\anaconda3\lib\site-packages (from thinc<8.3.0,>=8.2.2->spacy) (0.1.4)

Requirement already satisfied: colorama in c:\users\aravind\anaconda3\lib\site-packages (from tqdm<5.0.0,>=4.38.0->spacy) (0.4.6)

Requirement already satisfied: click<9.0.0,>=7.1.1 in c:\users\aravind\anaconda3\lib\site-packages (from typer<0.10.0,>=0.3.0->spacy) (8.0.4)

Requirement already satisfied: cloudpathlib<0.17.0,>=0.7.0 in c:\users\aravind\anaconda3\lib\site-packages (from weasel<0.4.0,>=0.1.0->spacy) (0.16.0)

Requirement already satisfied: MarkupSafe>=2.0 in c:\users\aravind\anaconda3\lib\site-packages (from jinja2->spacy) (2.1.1)

Requirement already satisfied: marisa-trie>=0.7.7 in c:\users\aravind\anaconda3\lib\site-packages (from language-data>=1.2->langcodes<4.0.0,>=3.2.0->spacy) (1.1.1)

Collecting en-core-web-sm==3.7.1

Using cached

[https://github.com/explosion/spacy-models/releases/download/en\\_core\\_web\\_sm-3.7.1/en\\_core\\_web\\_sm-3.7.1-py3-none-any.whl](https://github.com/explosion/spacy-models/releases/download/en_core_web_sm-3.7.1/en_core_web_sm-3.7.1-py3-none-any.whl) (12.8 MB)

Requirement already satisfied: spacy<3.8.0,>=3.7.2 in c:\users\aravind\anaconda3\lib\site-packages (from en-core-web-sm==3.7.1) (3.7.4)

Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (3.0.12)

Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (1.0.5)

Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (1.0.10)

Requirement already satisfied: cymem<2.1.0,>=2.0.2 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (2.0.8)

Requirement already satisfied: preshed<3.1.0,>=3.0.2 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (3.0.9)

Requirement already satisfied: thinc<8.3.0,>=8.2.2 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (8.2.3)

Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (1.1.2)

Requirement already satisfied: srsly<3.0.0,>=2.4.3 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (2.4.8)

Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (2.0.10)

Requirement already satisfied: weasel<0.4.0,>=0.1.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (0.3.4)

Requirement already satisfied: typer<0.10.0,>=0.3.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (0.9.4)

Requirement already satisfied: smart-open<7.0.0,>=5.2.1 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (5.2.1)

Requirement already satisfied: tqdm<5.0.0,>=4.38.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (4.65.0)

Requirement already satisfied: requests<3.0.0,>=2.13.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (2.31.0)

Requirement already satisfied: pydantic!=1.8,!1.8.1,<3.0.0,>=1.7.4 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (1.10.8)

Requirement already satisfied: jinja2 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (3.1.2)

Requirement already satisfied: setuptools in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (68.0.0)

Requirement already satisfied: packaging>=20.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (23.1)

Requirement already satisfied: langcodes<4.0.0,>=3.2.0 in c:\users\aravind\anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (3.4.0)

Requirement already satisfied: numpy>=1.19.0 in c:\users\aravind\

```

anaconda3\lib\site-packages (from spacy<3.8.0,>=3.7.2->en-core-web-
sm==3.7.1) (1.24.3)
Requirement already satisfied: language-data>=1.2 in c:\users\aravind\
anaconda3\lib\site-packages (from langcodes<4.0.0,>=3.2.0-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (1.2.0)
Requirement already satisfied: typing-extensions>=4.2.0 in c:\users\
aravind\anaconda3\lib\site-packages (from pydantic!=1.8,!
=1.8.1,<3.0.0,>=1.7.4->spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1)
(4.7.1)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\
aravind\anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\aravind\
anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\aravind\
anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (1.26.16)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\aravind\
anaconda3\lib\site-packages (from requests<3.0.0,>=2.13.0-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (2023.7.22)
Requirement already satisfied: blis<0.8.0,>=0.7.8 in c:\users\aravind\
anaconda3\lib\site-packages (from thinc<8.3.0,>=8.2.2-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (0.7.11)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in c:\users\
aravind\anaconda3\lib\site-packages (from thinc<8.3.0,>=8.2.2-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (0.1.4)
Requirement already satisfied: colorama in c:\users\aravind\anaconda3\
lib\site-packages (from tqdm<5.0.0,>=4.38.0->spacy<3.8.0,>=3.7.2->en-
core-web-sm==3.7.1) (0.4.6)
Requirement already satisfied: click<9.0.0,>=7.1.1 in c:\users\
aravind\anaconda3\lib\site-packages (from typer<0.10.0,>=0.3.0-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (8.0.4)
Requirement already satisfied: cloudpathlib<0.17.0,>=0.7.0 in c:\
users\aravind\anaconda3\lib\site-packages (from weasel<0.4.0,>=0.1.0-
>spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1) (0.16.0)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\aravind\
anaconda3\lib\site-packages (from jinja2->spacy<3.8.0,>=3.7.2->en-
core-web-sm==3.7.1) (2.1.1)
Requirement already satisfied: marisa-trie>=0.7.7 in c:\users\aravind\
anaconda3\lib\site-packages (from language-data>=1.2-
>langcodes<4.0.0,>=3.2.0->spacy<3.8.0,>=3.7.2->en-core-web-sm==3.7.1)
(1.1.1)
[+] Download and installation successful
You can now load the package via spacy.load('en_core_web_sm')

Requirement already satisfied: beautifulsoup4 in c:\users\aravind\
anaconda3\lib\site-packages (4.12.2)
Requirement already satisfied: soupsieve>1.2 in c:\users\aravind\
anaconda3\lib\site-packages (from beautifulsoup4) (2.4)

```

```
Requirement already satisfied: textblob in c:\users\aravind\anaconda3\lib\site-packages (0.18.0.post0)
Requirement already satisfied: nltk>=3.8 in c:\users\aravind\anaconda3\lib\site-packages (from textblob) (3.8.1)
Requirement already satisfied: click in c:\users\aravind\anaconda3\lib\site-packages (from nltk>=3.8->textblob) (8.0.4)
Requirement already satisfied: joblib in c:\users\aravind\anaconda3\lib\site-packages (from nltk>=3.8->textblob) (1.2.0)
Requirement already satisfied: regex>=2021.8.3 in c:\users\aravind\anaconda3\lib\site-packages (from nltk>=3.8->textblob) (2022.7.9)
Requirement already satisfied: tqdm in c:\users\aravind\anaconda3\lib\site-packages (from nltk>=3.8->textblob) (4.65.0)
Requirement already satisfied: colorama in c:\users\aravind\anaconda3\lib\site-packages (from click->nltk>=3.8->textblob) (0.4.6)
```

```
import warnings
warnings.filterwarnings('ignore')
```

```
import pandas as pd
```

```
#READ DATASET
```

```
df = pd.read_csv('https://raw.githubusercontent.com/laxmimerit/All-CSV-ML-Data-Files-Download/master/twitter_sentiment.csv',header=None,index_col = 0)
df.head()
```

```

      1      2
3
0
2401 Borderlands Positive im getting on borderlands and i will murder yo...
2401 Borderlands Positive I am coming to the borders and I will kill you...
2401 Borderlands Positive im getting on borderlands and i will kill you ...
2401 Borderlands Positive im coming on borderlands and i will murder you...
2401 Borderlands Positive im getting on borderlands 2 and i will murder ...
```

```
df = df[[2,3]].reset_index(drop= True)
```

```
df.columns = ['sentiment','text']
df.head()
```

```

  sentiment      text
0  Positive  im getting on borderlands and i will murder yo...
1  Positive  I am coming to the borders and I will kill you...
2  Positive  im getting on borderlands and i will kill you ...
```

```
3 Positive im coming on borderlands and i will murder you...
4 Positive im getting on borderlands 2 and i will murder ...
```

### #BASIC ANALYSIS

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 75682 entries, 0 to 75681
Data columns (total 2 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   sentiment    75682 non-null  object
1   text         74996 non-null  object
dtypes: object(2)
memory usage: 1.2+ MB
```

```
df.isnull().sum()
df.dropna(inplace= True)
```

```
sum(df['text'].apply(len)>5),sum(df['text'].apply(len)<=5)

(72883, 2113)
```

```
print(df.shape)
df=df[df['text'].apply(len)>5]
print(df.shape)
```

```
(74996, 2)
(72883, 2)
```

```
df['sentiment'].value_counts()
```

```
sentiment
Negative      22020
Positive      20186
Neutral       17899
Irrelevant    12778
Name: count, dtype: int64
```

## prerocessing

```
!pip install
git+https://github.com/laxmimerit/preprocess_kgptalkie.git --upgrade
--force-reinstall
```

```
Collecting git+https://github.com/laxmimerit/preprocess_kgptalkie.git
  Cloning https://github.com/laxmimerit/preprocess_kgptalkie.git to
  c:\users\aravind\appdata\local\temp\pip-req-build-e6bh26hw
  Resolved https://github.com/laxmimerit/preprocess_kgptalkie.git to
```

```

commit 96bf02872d9756f29d6cddb8aafaedcd2a39bbb4
  Preparing metadata (setup.py): started
  Preparing metadata (setup.py): finished with status 'done'
Building wheels for collected packages: preprocess-kgptalkie
  Building wheel for preprocess-kgptalkie (setup.py): started
  Building wheel for preprocess-kgptalkie (setup.py): finished with
status 'done'
  Created wheel for preprocess-kgptalkie:
filename=preprocess_kgptalkie-0.1.3-py3-none-any.whl size=7667
sha256=70b87e0fdd205345a53aab4a1adfb86735def6b2661910220b3966d2c8180bd
4
  Stored in directory: C:\Users\Aravind\AppData\Local\Temp\pip-ephem-
wheel-cache-eg2e0qtg\wheels\74\fe\
05\0d013c54ae5e4afb77c6c480378063827c84f57ea5f554e072
Successfully built preprocess-kgptalkie
Installing collected packages: preprocess-kgptalkie
  Attempting uninstall: preprocess-kgptalkie
    Found existing installation: preprocess-kgptalkie 0.1.3
    Uninstalling preprocess-kgptalkie-0.1.3:
      Successfully uninstalled preprocess-kgptalkie-0.1.3
Successfully installed preprocess-kgptalkie-0.1.3

  Running command git clone --filter=blob:none --quiet
https://github.com/laxmimerit/preprocess_kgptalkie.git 'C:\Users\
Aravind\AppData\Local\Temp\pip-req-build-e6bh26hw'

import preprocess_kgptalkie as ps
df.columns

Index(['sentiment', 'text', dtype='object'])

df = ps.get_basic_features(df)

df.columns

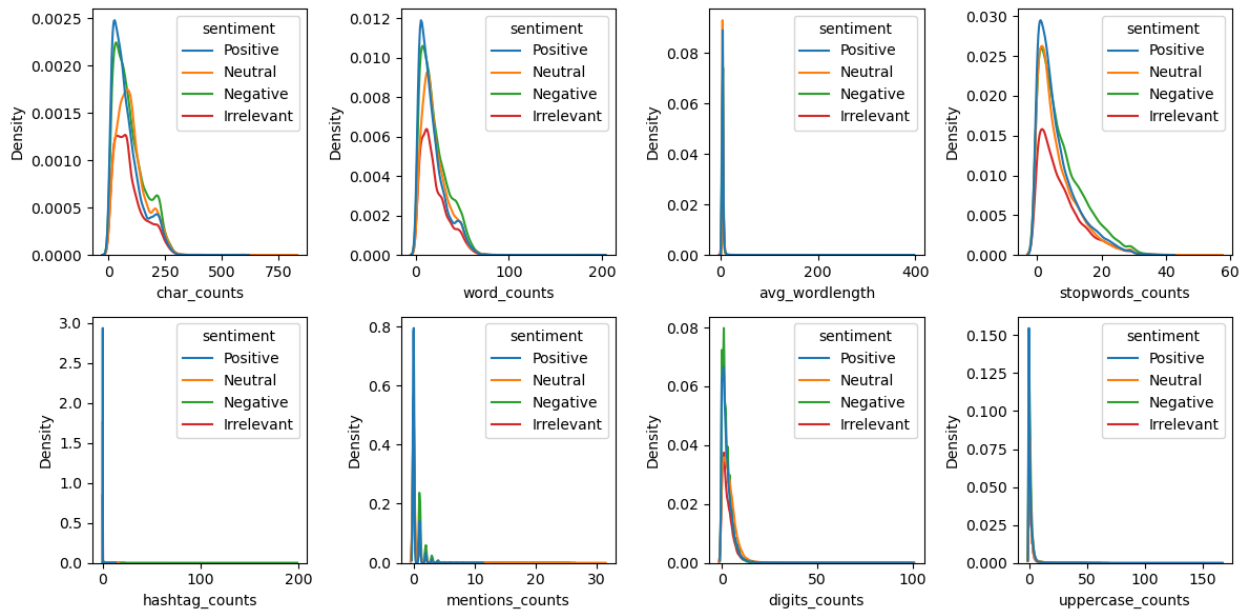
Index(['sentiment', 'text', 'char_counts', 'word_counts',
'avg_wordlength',
      'stopwords_counts', 'hashtag_counts', 'mentions_counts',
      'digits_counts', 'uppercase_counts'],
      dtype='object')

df.head()
import matplotlib.pyplot as plt
import seaborn as sns
plt.figure(figsize=(12,6))
num_cols=df.select_dtypes(include='number').columns

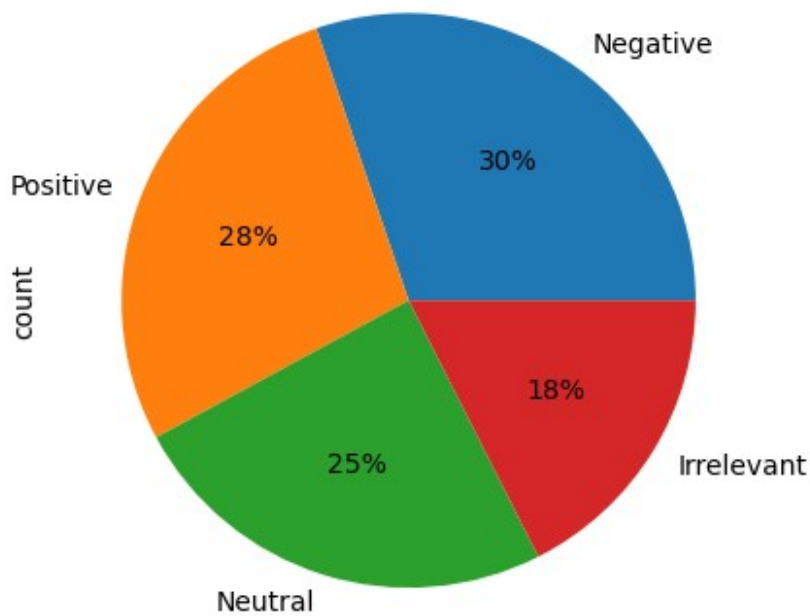
for index ,col in enumerate(num_cols):
    plt.subplot(2,4 ,index+1)
    sns.kdeplot(data = df,x=col,hue="sentiment",fill= False)

```

```
plt.tight_layout()
plt.show()
```



```
df['sentiment'].value_counts().plot(kind='pie', autopct="%1.0f%%")
<Axes: ylabel='count'>
```





# wordcloud analysis

```
!pip install wordcloud
```

```
Requirement already satisfied: wordcloud in c:\users\aravind\
anaconda3\lib\site-packages (1.9.3)
Requirement already satisfied: numpy>=1.6.1 in c:\users\aravind\
anaconda3\lib\site-packages (from wordcloud) (1.24.3)
Requirement already satisfied: pillow in c:\users\aravind\anaconda3\
lib\site-packages (from wordcloud) (9.4.0)
Requirement already satisfied: matplotlib in c:\users\aravind\
anaconda3\lib\site-packages (from wordcloud) (3.7.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\aravind\
anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\aravind\
anaconda3\lib\site-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\aravind\
anaconda3\lib\site-packages (from matplotlib->wordcloud) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\aravind\
anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\aravind\
anaconda3\lib\site-packages (from matplotlib->wordcloud) (23.1)
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in c:\users\
aravind\anaconda3\lib\site-packages (from matplotlib->wordcloud)
(3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\
aravind\anaconda3\lib\site-packages (from matplotlib->wordcloud)
(2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\aravind\anaconda3\
lib\site-packages (from python-dateutil>=2.7->matplotlib->wordcloud)
(1.16.0)
```

```
from wordcloud import WordCloud, STOPWORDS
stopwords = set(STOPWORDS)
```

```
stopwords
```

```
{'a',
 'about',
 'above',
 'after',
 'again',
 'against',
 'all',
 'also',
 'am',
 'an',
 'and',
 'any',
 'are',
```

"aren't",  
'as',  
'at',  
'be',  
'because',  
'been',  
'before',  
'being',  
'below',  
'between',  
'both',  
'but',  
'by',  
'can',  
"can't",  
'cannot',  
'com',  
'could',  
"couldn't",  
'did',  
"didn't",  
'do',  
'does',  
"doesn't",  
'doing',  
"don't",  
'down',  
'during',  
'each',  
'else',  
'ever',  
'few',  
'for',  
'from',  
'further',  
'get',  
'had',  
"hadn't",  
'has',  
"hasn't",  
'have',  
"haven't",  
'having',  
'he',  
"he'd",  
"he'll",  
"he's",  
'hence',  
'her',

'here',  
"here's",  
'hers',  
'herself',  
'him',  
'himself',  
'his',  
'how',  
"how's",  
'however',  
'http',  
'i',  
"i'd",  
"i'll",  
"i'm",  
"i've",  
'if',  
'in',  
'into',  
'is',  
"isn't",  
'it',  
"it's",  
'its',  
'itself',  
'just',  
'k',  
"let's",  
'like',  
'me',  
'more',  
'most',  
"mustn't",  
'my',  
'myself',  
'no',  
'nor',  
'not',  
'of',  
'off',  
'on',  
'once',  
'only',  
'or',  
'other',  
'otherwise',  
'ought',  
'our',  
'ours',

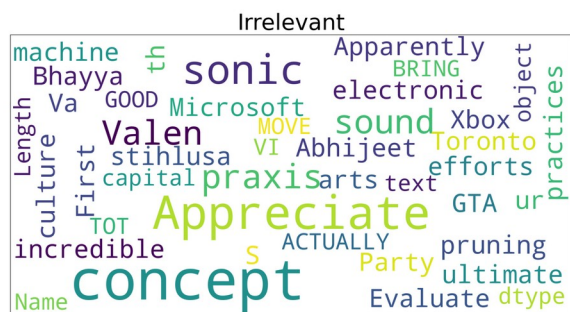
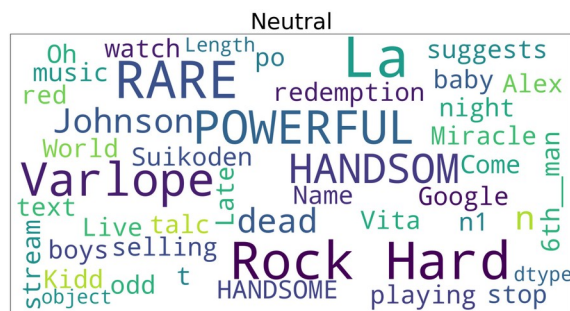
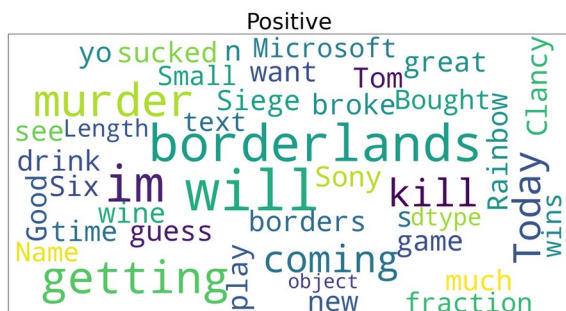
'ourselves',  
'out',  
'over',  
'own',  
'r',  
'same',  
'shall',  
"shan't",  
'she',  
"she'd",  
"she'll",  
"she's",  
'should',  
"shouldn't",  
'since',  
'so',  
'some',  
'such',  
'than',  
'that',  
"that's",  
'the',  
'their',  
'theirs',  
'them',  
'themselves',  
'then',  
'there',  
"there's",  
'therefore',  
'these',  
'they',  
"they'd",  
"they'll",  
"they're",  
"they've",  
'this',  
'those',  
'through',  
'to',  
'too',  
'under',  
'until',  
'up',  
'very',  
'was',  
"wasn't",  
'we',  
"we'd",

```
"we'll",
"we're",
"we've",
'were',
"weren't",
'what',
"what's",
'when',
"when's",
'where',
"where's",
'which',
'while',
'who',
"who's",
'whom',
'why',
"why's",
'with',
"won't",
'would',
"wouldn't",
'www',
'you',
"you'd",
"you'll",
"you're",
"you've",
'your',
'yours',
'yourself',
'yourselves'}
```

```
wordcloud = WordCloud(background_color =
'white',stopwords=stopwords,max_words=300,max_font_size=40,scale=5).ge
nerate(str(df['text']))
plt.imshow(wordcloud)
```

```
<matplotlib.image.AxesImage at 0x2c58e884310>
```





## DATA CLEANING

```
df['text'] = df['text'].apply(lambda x: x.lower())
df['text'] = df['text'].apply(lambda x: ps.remove_urls(x))
df['text'] = df['text'].apply(lambda x: ps.remove_html_tags(x))
df['text'] = df['text'].apply(lambda x: ps.remove_special_chars(x))
df['text'] = df['text'].apply(lambda x: ps.remove_rt(x))
```

## train test split

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(df['text'],
df['sentiment'], test_size=0.2, random_state=0)

X_train.shape,X_test.shape

((58306,), (14577,))
```

## MODEL BUILDING AND TESTING

```
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.ensemble import RandomForestClassifier
from sklearn.pipeline import Pipeline
```

```

clf = Pipeline([('tfidf', TfidfVectorizer()), ('clf',
RandomForestClassifier(n_estimators=100, n_jobs=-1))])
clf.fit(X_train, y_train)

Pipeline(steps=[('tfidf', TfidfVectorizer()),
                 ('clf', RandomForestClassifier(n_jobs=-1))])

```

*# evaluation*

```
from sklearn.metrics import classification_report
```

```
Y_pred = clf.predict(X_test)
```

```
print( classification_report
(y_test,Y_pred))
```

	precision	recall	f1-score	support
Irrelevant	0.98	0.86	0.91	2561
Negative	0.92	0.96	0.94	4481
Neutral	0.92	0.92	0.92	3553
Positive	0.91	0.95	0.93	3982
accuracy			0.93	14577
macro avg	0.93	0.92	0.92	14577
weighted avg	0.93	0.93	0.93	14577

```
from sklearn.metrics import accuracy_score
```

```
predictions = clf.predict(X_test)
print(accuracy_score(y_test, predictions))
```

```
0.9266652946422447
```

```
import pickle
```

```
pickle.dump(clf, open('Twitter_sentiment.pkl', 'wb'))
```

```
X_test.tolist()[20]
```

```

['thank god',
 'i dont know why a weird ass district using this wack ass software
when we couldve used google classroom',
 '2ksupport nba2k my boy whos new to the game and relies on me to
teach him how to play was naughty when creating a ticket to figure out
why he left a reply a few time claiming he was banned for boosting
again he doesnt even know how to play other than what i show him 12',
 'tristate got the best video we international are in this scene',
 'callofduty is currently an absolute mess',

```



'revenge',  
'dell 1 xps 17 9700 graphics review card missing and flickering  
screen via r dellxps ift tt 2ps3pbw about a whole week straight ago i  
personally was just playing a console game and the computer was  
suddenly blue screened with a fatal error indicating a crashed gpu  
upon rebooting maybe the nvidia gtx was missing link in device',  
'i havent played hearthstone for years but my god this card is as  
cute as heck picfacebookcom 7lqtnrndgl',  
'jnjcares good luck hope sales go before profit available for  
everyone',  
'everything i like about call of duty doesnt apply in mw just keep  
the simple shit in and ppl would play it',  
'poor nintendo',  
'simple janie clever janie she is the good sister who never causes  
problems meet her now in janies secrets',  
'pubgmobile pubg pubgmobile \_ in pubg \_ support update my pubg app  
after logging in i am unreachable pictwittercom voundlabqy',  
'all of saying these videos are getting me really pretty excited for  
lor i doubt it would actually ever fully replace those hearthstone  
forms but most videos like this show tell you how fun the game can  
be',  
'twittercom minteractapp s https tco ubdtre9iud',  
'you ultimate pruning machine stihlusa gta series has every 8 bar and  
chain in size making the perfect for landscapers or homeowners  
shopping for a smaller machine vs manually gardening with a  
instagramcompb74nelb3eg',  
'thank you gachamber in your leadership',  
'riotgames leagueoflegends whoever i should tag but can you actually  
get to banning actual toxic frustrating players which is of it that  
too hard to actually do you guys are a million dollar gaming company  
worldwide and all yet you cant manage you to do that something about  
hate the vile players that make this game unpleasant',  
'and also bought borderlands on the switch we love it that series and  
cant wait to play on the second go',  
'ups 2 bad she is not n place wen ups cust reps decided 2oss my 1600  
last nov 2019 and ur reps wont go something 2 a oid r ur rand  
integrity s called n2 question trust is a lot 2 mo 2 disability  
benefit']  
  
clf.predict(['i have done'])  
  
array(['Positive'], dtype=object)