```
# This Python 3 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/) that gets preserved as output when you create a version using "Sav
# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the current session
```

⇥ /kaggle/input/loksabha-election-2024-twitter-data/LokSabha_Election_2024_Tweets.csv


```
!pip install nltk
```

⇥ Requirement already satisfied: nltk in /usr/local/lib/python3.10/dist-packages (3.8.1)
   Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk) (8.1.7)
   Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk) (1.4.2)
   Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk) (2024.9.11)
   Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk) (4.66.5)


```
!pip install wordcloud
```

⇥ Requirement already satisfied: wordcloud in /opt/conda/lib/python3.10/site-packages (1.9.3)
   Requirement already satisfied: numpy>=1.6.1 in /opt/conda/lib/python3.10/site-packages (from wordcloud) (1.26.4)
   Requirement already satisfied: pillow in /opt/conda/lib/python3.10/site-packages (from wordcloud) (10.3.0)
   Requirement already satisfied: matplotlib in /opt/conda/lib/python3.10/site-packages (from wordcloud) (3.7.5)
   Requirement already satisfied: contourpy>=1.0.1 in /opt/conda/lib/python3.10/site-packages (from matplotlib->wordcloud) (1.2.1)
   Requirement already satisfied: cycler>=0.10 in /opt/conda/lib/python3.10/site-packages (from matplotlib->wordcloud) (0.12.1)
   Requirement already satisfied: fonttools>=4.22.0 in /opt/conda/lib/python3.10/site-packages (from matplotlib->wordcloud) (4.53.0)
   Requirement already satisfied: kiwisolver>=1.0.1 in /opt/conda/lib/python3.10/site-packages (from matplotlib->wordcloud) (1.4.5)
   Requirement already satisfied: packaging>=20.0 in /opt/conda/lib/python3.10/site-packages (from matplotlib->wordcloud) (21.3)
   Requirement already satisfied: pyparsing>=2.3.1 in /opt/conda/lib/python3.10/site-packages (from matplotlib->wordcloud) (3.1.2)
   Requirement already satisfied: python-dateutil>=2.7 in /opt/conda/lib/python3.10/site-packages (from matplotlib->wordcloud) (2.9.0.p
   Requirement already satisfied: six>=1.5 in /opt/conda/lib/python3.10/site-packages (from python-dateutil>=2.7->matplotlib->wordcloud

◄   ►


```
!pip install textblob
```

⇥ Requirement already satisfied: textblob in /usr/local/lib/python3.10/dist-packages (0.17.1)
   Requirement already satisfied: nltk>=3.1 in /usr/local/lib/python3.10/dist-packages (from textblob) (3.8.1)
   Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from nltk>=3.1->textblob) (8.1.7)
   Requirement already satisfied: joblib in /usr/local/lib/python3.10/dist-packages (from nltk>=3.1->textblob) (1.4.2)
   Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.10/dist-packages (from nltk>=3.1->textblob) (2024.9.11)
   Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from nltk>=3.1->textblob) (4.66.5)


```
# Import Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import plotly.express as px

# Libraries for Sentiment Analysis
import re
import nltk
from nltk.corpus import stopwords
from nltk.corpus import wordnet
from nltk.stem import WordNetLemmatizer
from textblob import TextBlob
from wordcloud import WordCloud

# to avoid warnings
import warnings
warnings.filterwarnings('ignore')
```

IMPORTING THE DATA


```
lkele=pd.read_csv("/content/LokSabha_Election_2024_Tweets.csv")
```

```
lkele["text"]
```

	text
0	South India 2024 Loksabha Opinion Poll 132 sea...
1	Veteran actor Nana Patekar says, "There is no ...
2	South India BJP 2019 Loksabha Karnataka 25 Tel...
3	Issh baar 357 seats in Loksabha election.
4	Total number of times MMS won his own lok sabh...
...	...
995	He will make sure Congress puts a dummy candid...
996	This ll reverse, because TMC is facing huge an...
997	Sabseh kathin yatra I.N.D.I. ALLIANCE meh Akhi...
998	Bhai tumhari inhi harkato ki vjh se Pilot saab...
999	Very well said, when a leader couldnt win an e...

1000 rows × 1 columns

```
print(lkele.columns)
lkele.shape
```

```
Index(['link', 'text', 'date', 'No_of_likes', 'No_of_comments'], dtype='object')
(1000, 5)
```

```
lkele.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 5 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   link            1000 non-null   object
 1   text            995 non-null    object
 2   date            1000 non-null   object
 3   No_of_likes     1000 non-null   int64
 4   No_of_comments  1000 non-null   int64
dtypes: int64(2), object(3)
memory usage: 39.2+ KB
```

DATA PREPROCESSING

```
lkele.dropna(inplace=True)
lkele.shape
```

```
(995, 5)
```

```
mean_like= lkele['No_of_likes'].value_counts().mean()
print(mean_like)
```

```
3.618181818181818
```

Exploratory Data Analysis

```
#To know the trend of number oof tweets along the time

lkele["date"]=pd.to_datetime(lkele["date"],format='%b %d, %Y · %I:%M %p UTC')
lkele["date_only"]=lkele["date"].dt.date
fig=px.histogram(lkele,x="date_only")
fig.show()
```

SENTIMENT ANALYSIS

```python
import nltk
nltk.download('wordnet')
nltk.download('stopwords')  # If you're using stopwords as well
```

```
[nltk_data] Downloading package wordnet to /root/nltk_data...
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]    Unzipping corpora/stopwords.zip.
True
```

```python
def clean(text):

    # Convert text to lowercase
    text = text.lower()

    # Replace anything other than alphabets a-z with a space
    ext = re.sub('[^a-z]', ' ', text)

    # Split the text into single words
    text = text.split()

    # Initialize WordNetLemmatizer
    lm = WordNetLemmatizer()

    # Lemmatize words and remove stopwords
    text = [lm.lemmatize(word) for word in text if word not in set(
        stopwords.words('english'))]

    # Join the words back into a sentence
    text = ' '.join(word for word in text)

    return text
```

```python
def getpolarity(text):
    return TextBlob(text).sentiment.polarity

def getsubjectivity(text):
    return TextBlob(text).sentiment.subjectivity

def getAnalysis(score):
    if score < 0:
        return 'negative'
    elif score == 0:
        return 'neutral'
    else:
        return 'positive'
```

```
lkele["cleantxt"]=lkele["text"].apply(clean)
print(lkele.head())
```

```
                                            link  \
0  https://twitter.com/Politics_2022_/status/1739...
1  https://twitter.com/LokmatTimes_ngp/status/173...
2  https://twitter.com/Politics_2022_/status/1739...
3  https://twitter.com/rohanayak/status/173926994...
4  https://twitter.com/BellamSwathi/status/173919...

                                            text              date  \
0  South India 2024 Loksabha Opinion Poll 132 sea... 2023-12-25 12:05:00
1  Veteran actor Nana Patekar says, "There is no ... 2023-12-25 13:00:00
2  South India BJP 2019 Loksabha Karnataka 25 Tel... 2023-12-25 12:21:00
3         Issh baar 357 seats in Loksabha election. 2023-12-25 13:00:00
4  Total number of times MMS won his own lok sabh... 2023-12-25 08:19:00

   No_of_likes  No_of_comments   date_only  \
0          324              27  2023-12-25
1            0               0  2023-12-25
2          142              18  2023-12-25
3            0               0  2023-12-25
4          589              80  2023-12-25

                                         cleantxt
0  south india 2024 loksabha opinion poll 132 sea...
1  veteran actor nana patekar says, "there altern...
2  south india bjp 2019 loksabha karnataka 25 tel...
3            issh baar 357 seat loksabha election.
4  total number time mm lok sabha seat - zero hen...
```

```
lkele["polarity"]=lkele["cleantxt"].apply(getpolarity)
```

```
lkele["subjectivity"]=lkele["cleantxt"].apply(getsubjectivity)
```

```
lkele["analysis"]=lkele["polarity"].apply(getAnalysis)
```

```
lkele.head()
```

	link	text	date	No_of_likes	No_of_comments	date_only	cleantxt	polarity	subjecti
0	https://twitter.com/Politics_2022_/status/1739...	South India 2024 Loksabha Opinion Poll 132 sea...	2023-12-25 12:05:00	324	27	2023-12-25	south india 2024 loksabha opinion poll 132 sea...	0.800000	0.40

```
# how much data is positive/negetive/neutral
plt.style.use('dark_background')  # Adding black theme

# Define colors for each bar
colors = ['orange', 'blue', 'red']

plt.figure(figsize=(7, 5))
(lkele.analysis.value_counts(normalize=True) * 100).plot.bar(color=colors)
plt.ylabel("%age of tweets")
plt.title("Loksabha Election Polarity")
plt.show()
```

```python
from wordcloud import WordCloud, STOPWORDS

def word_cloud(wd_list):
    stopwords = set(STOPWORDS)
    all_words = ' '.join(wd_list)
    wordcloud = WordCloud(background_color='black',
                          stopwords=stopwords,
                          width=1600, height=800, max_words=100, max_font_size=200,
                          colormap="viridis").generate(all_words)
    plt.figure(figsize=(12, 10))
    plt.axis('off')
    plt.imshow(wordcloud)

word_cloud(lkele['cleantxt'][:5000])

# This code is modified by Susobhan Akhuli
```



```python
congress_tweets = lkele[lkele['cleantxt'].str.contains('congress')]

# Filter tweets mentioning 'bjp'
bjp_tweets = lkele[lkele['cleantxt'].str.contains('bjp')]
```

```python
congress_tweets.head()
```

	link	text	date	No_of_likes	No_of_comments	date_only	cleantxt	polarity	sub
2	https://twitter.com/Politics_2022_/status/1739...	South India BJP 2019 Loksabha Karnataka 25 Tel...	2023-12-25 12:21:00	142	18	2023-12-25	south india bjp 2019 loksabha karnataka 25 tel...	-0.300000	
		Yatnal had					yatnal talk		

```
congress_tweets.shape
```

```
(154, 10)
```

```
bjp_tweets.shape
```
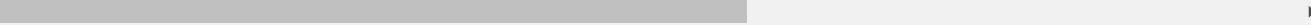
```
(281, 10)
```

```
bjp_tweets.head()
```

	link	text	date	No_of_likes	No_of_comments	date_only	cleantxt	polarity	subject
0	https://twitter.com/Politics_2022_/status/1739...	South India 2024 Loksabha Opinion Poll 132 sea...	2023-12-25 12:05:00	324	27	2023-12-25	south india 2024 loksabha opinion poll 132 sea...	0.800000	0.4

```python
# how much data is positive/negetive/neutral
plt.style.use('dark_background')  # Adding black theme

# Define colors for each bar
colors = ['orange', 'blue', 'red']

plt.figure(figsize=(7, 5))
(congress_tweets.analysis.value_counts(normalize=True) * 100).plot.bar(color=colors)
plt.ylabel("%age of tweets")
plt.title("Congress' Polarity")
plt.show()
```
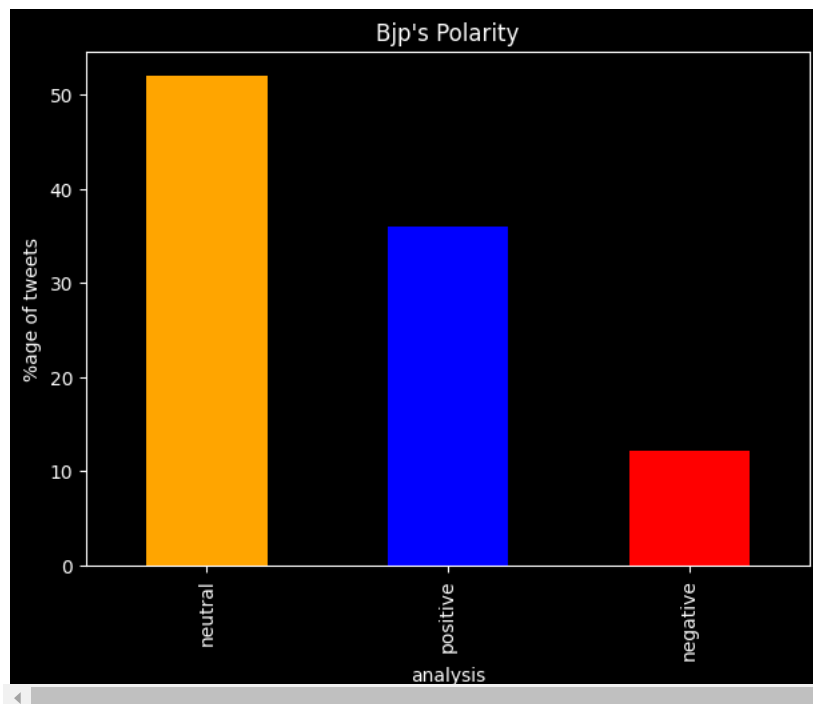
Congress' Polarity

```python
# how much data is positive/negetive/neutral
plt.style.use('dark_background')  # Adding black theme

# Define colors for each bar
colors = ['orange', 'blue', 'red']

plt.figure(figsize=(7, 5))
(bjp_tweets.analysis.value_counts(normalize=True) * 100).plot.bar(color=colors)
plt.ylabel("%age of tweets")
plt.title("Bjp's Polarity")
plt.show()
```



```python
namo_tweets = lkele[lkele['cleantxt'].str.contains('narendramodi')]
namo_tweets = lkele[lkele['cleantxt'].str.contains('modi')]
namo_tweets = lkele[lkele['cleantxt'].str.contains('namo')]
namo_tweets = lkele[lkele['cleantxt'].str.contains('pm')]
namo_tweets = lkele[lkele['cleantxt'].str.contains('primeminister')]


# Filter tweets mentioning 'bjp'
pappu_tweets = lkele[lkele['cleantxt'].str.contains('raghul')]
```