

Project on Twitter analysis Using Link by MANISH SINGH THENUA

In [1]: # Import the Libraries

```
import tweepy
from textblob import TextBlob
from wordcloud import WordCloud
import pandas as pd
import numpy as np
import re
import matplotlib.pyplot as plt
```

In [2]: !pip install tweepy

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: tweepy in c:\users\user\appdata\roaming\python
\python39\site-packages (4.13.0)
Requirement already satisfied: requests-oauthlib<2,>=1.2.0 in c:\users\user\appdata\roaming\python\python39\site-packages (from tweepy) (1.3.1)
Requirement already satisfied: oauthlib<4,>=3.2.0 in c:\users\user\appdata\roaming\python\python39\site-packages (from tweepy) (3.2.2)
Requirement already satisfied: requests<3,>=2.27.0 in c:\users\user\appdata\roaming\python\python39\site-packages (from tweepy) (2.28.2)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\programdata\anaconda3\lib\site-packages (from requests<3,>=2.27.0->tweepy) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\programdata\anaconda3\lib\site-packages (from requests<3,>=2.27.0->tweepy) (2022.9.14)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\user\appdata\roaming\python\python39\site-packages (from requests<3,>=2.27.0->tweepy) (1.22)
Requirement already satisfied: idna<4,>=2.5 in c:\users\user\appdata\roaming\python\python39\site-packages (from requests<3,>=2.27.0->tweepy) (2.6)
```

In [3]: !pip install textblob

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: textblob in c:\users\user\appdata\roaming\python\python39\site-packages (0.17.1)
Requirement already satisfied: nltk>=3.1 in c:\programdata\anaconda3\lib\site-packages (from textblob) (3.7)
Requirement already satisfied: click in c:\programdata\anaconda3\lib\site-packages (from nltk>=3.1->textblob) (8.0.4)
Requirement already satisfied: regex>=2021.8.3 in c:\programdata\anaconda3\lib\site-packages (from nltk>=3.1->textblob) (2022.7.9)
Requirement already satisfied: joblib in c:\users\user\appdata\roaming\python\python39\site-packages (from nltk>=3.1->textblob) (1.2.0)
Requirement already satisfied: tqdm in c:\programdata\anaconda3\lib\site-packages (from nltk>=3.1->textblob) (4.64.1)
Requirement already satisfied: colorama in c:\users\user\appdata\roaming\python\python39\site-packages (from click->nltk>=3.1->textblob) (0.4.6)
```

In [4]: !pip install wordcloud

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: wordcloud in c:\users\user\appdata\roaming\python\python39\site-packages (1.8.2.2)
Requirement already satisfied: matplotlib in c:\users\user\appdata\roaming\python\python39\site-packages (from wordcloud) (3.6.3)
Requirement already satisfied: pillow in c:\programdata\anaconda3\lib\site-packages (from wordcloud) (9.2.0)
Requirement already satisfied: numpy>=1.6.1 in c:\programdata\anaconda3\lib\site-packages (from wordcloud) (1.21.5)
Requirement already satisfied: fonttools>=4.22.0 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (4.25.0)
Requirement already satisfied: python-dateutil>=2.7 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (2.8.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\user\appdata\roaming\python\python39\site-packages (from matplotlib->wordcloud) (1.0.7)
Requirement already satisfied: packaging>=20.0 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (21.3)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.4.2)
Requirement already satisfied: cycler>=0.10 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: pyparsing>=2.2.1 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (3.0.9)
Requirement already satisfied: six>=1.5 in c:\programdata\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib->wordcloud) (1.16.0)
```

In [5]: !pip install snscreape

```
Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: snscreape in c:\users\user\appdata\roaming\python\python39\site-packages (0.6.2.20230320)
Requirement already satisfied: requests[socks] in c:\users\user\appdata\roaming\python\python39\site-packages (from snscreape) (2.28.2)
Requirement already satisfied: lxml in c:\programdata\anaconda3\lib\site-packages (from snscreape) (4.9.1)
Requirement already satisfied: filelock in c:\programdata\anaconda3\lib\site-packages (from snscreape) (3.6.0)
Requirement already satisfied: beautifulsoup4 in c:\programdata\anaconda3\lib\site-packages (from snscreape) (4.11.1)
Requirement already satisfied: soupsieve>1.2 in c:\programdata\anaconda3\lib\site-packages (from beautifulsoup4->snscreape) (2.3.1)
Requirement already satisfied: idna<4,>=2.5 in c:\users\user\appdata\roaming\python\python39\site-packages (from requests[socks]->snscreape) (2.6)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\programdata\anaconda3\lib\site-packages (from requests[socks]->snscreape) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in c:\programdata\anaconda3\lib\site-packages (from requests[socks]->snscreape) (2022.9.14)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\user\appdata\roaming\python\python39\site-packages (from requests[socks]->snscreape) (1.22)
Requirement already satisfied: PySocks!=1.5.7,>=1.5.6 in c:\programdata\anaconda3\lib\site-packages (from requests[socks]->snscreape) (1.7.1)
```

Tweet Analysis with Link of twitter account

```
In [6]: import sns_scrape.modules.twitter as sntwitter
import pandas as pd

query = "(from:elonmusk) until:2023-03-31 since:2022-01-01"
tweets = []
limit = 5000

for tweet in sntwitter.TwitterSearchScraper(query).get_items():

    if len(tweets) == limit:
        break
    else:
        tweets.append([tweet.date, tweet.user.username, tweet.content])

df = pd.DataFrame(tweets, columns=['Date', 'User', 'Tweet'])
print(df)
```

C:\Users\user\AppData\Local\Temp\ipykernel_9116\178922417.py:13: DeprecationWarning: content is deprecated, use rawContent instead
 tweets.append([tweet.date, tweet.user.username, tweet.content])

	Date	User	\
0	2023-03-30 23:59:08+00:00	elonmusk	
1	2023-03-30 22:02:28+00:00	elonmusk	
2	2023-03-30 07:44:08+00:00	elonmusk	
3	2023-03-30 07:15:48+00:00	elonmusk	
4	2023-03-30 07:13:10+00:00	elonmusk	
...	
4995	2022-07-19 15:59:19+00:00	elonmusk	
4996	2022-07-19 15:55:22+00:00	elonmusk	
4997	2022-07-19 07:00:38+00:00	elonmusk	
4998	2022-07-19 05:32:27+00:00	elonmusk	
4999	2022-07-19 02:40:08+00:00	elonmusk	

	Tweet
0	@mtaibbi 🎉
1	@iamharaldur !
2	@Galaxygamerone Coming soon
3	@Tesmanian_com It will per vehicle
4	@TPostMillennial !!
...	...
4995	To think that Fremont was just a derelict buil...
4996	Congrats to Tesla Fremont + Giga Nevada for ma...
4997	Please vote your shares with Tesla!
4998	@WholeMarsBlog Yeah
4999	@SawyerMerritt @SpaceX It's a start. About 100...

[5000 rows x 3 columns]

In [7]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5000 entries, 0 to 4999
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype  
---  -- 
 0   Date     5000 non-null    datetime64[ns, UTC]
 1   User     5000 non-null    object  
 2   Tweet    5000 non-null    object  
dtypes: datetime64[ns, UTC](1), object(2)
memory usage: 117.3+ KB
```

In [8]: `df.columns`

Out[8]: `Index(['Date', 'User', 'Tweet'], dtype='object')`

In [9]: `def remove_pattern(input_txt,pattern):
 r = re.findall(pattern, input_txt)
 for word in r:
 input_txt = re.sub(word,"",input_txt)
 return input_txt`

```
In [10]: # remove Pattern (@user)
df['clean_tweet'] = np.vectorize(remove_pattern)(df['Tweet'], "@[\w]*")
df.head(20)
```

Out[10]:

	Date	User	Tweet	clean_tweet
0	2023-03-30 23:59:08+00:00	elonmusk	@mtaibbi 🚀	🚀
1	2023-03-30 22:02:28+00:00	elonmusk	@iamharaldur !	!
2	2023-03-30 07:44:08+00:00	elonmusk	@Galaxygamerone Coming soon	Coming soon
3	2023-03-30 07:15:48+00:00	elonmusk	@Tesmanian_com It will per vehicle	It will per vehicle
4	2023-03-30 07:13:10+00:00	elonmusk	@TPostMillennial !!	!!
5	2023-03-30 06:37:44+00:00	elonmusk	Old joke about agnostic technologists building...	Old joke about agnostic technologists building...
6	2023-03-29 23:18:40+00:00	elonmusk	@BillyM2k Makes me sad	Makes me sad
7	2023-03-29 22:53:09+00:00	elonmusk	@GRDector Combined with excess government spen...	Combined with excess government spending, whi...
8	2023-03-29 22:51:48+00:00	elonmusk	@GRDector Serious issue. US policy has been to...	Serious issue. US policy has been too heavy-h...
9	2023-03-29 22:41:45+00:00	elonmusk	@boringcompany Making progress!	Making progress!
10	2023-03-29 21:50:32+00:00	elonmusk	@andst7 🚀	🚀
11	2023-03-29 21:43:57+00:00	elonmusk	@EvaFoxU 6 years ago	6 years ago
12	2023-03-29 19:06:50+00:00	elonmusk	@iheartmindy This deserves close scrutiny 😊	This deserves close scrutiny 😊
13	2023-03-29 19:05:37+00:00	elonmusk	@BillyM2k AI optimizing for the greatest under...	AI optimizing for the greatest understanding ...
14	2023-03-29 18:31:33+00:00	elonmusk	@Jason Yeah	Yeah
15	2023-03-29 18:02:34+00:00	elonmusk	@TwitterSafety Given that there are ~500 milli...	Given that there are ~500 million tweets per ...
16	2023-03-29 17:36:31+00:00	elonmusk	@TheChiefNerd Leading AGI developers will not ...	Leading AGI developers will not heed this war...
17	2023-03-29 07:21:39+00:00	elonmusk	@TeslaDiva99 Tesla car temp is automatically k...	Tesla car temp is automatically kept within a...
18	2023-03-29 07:15:58+00:00	elonmusk	@george_mack I thought both SpaceX and Tesla ...	I thought both SpaceX and Tesla each had less...
19	2023-03-28 19:56:26+00:00	elonmusk	@alx Sick burn lol	Sick burn lol

```
In [11]: # Remove special Characterstic number and punctuation  
df['clean_tweet'] = df['Tweet'].str.replace("[^a-zA-Z#]", " ")  
df.head(20)
```

C:\Users\user\AppData\Local\Temp\ipykernel_9116\2840319962.py:2: FutureWarning:
g: The default value of regex will change from True to False in a future vers
ion.

```
df['clean_tweet'] = df['Tweet'].str.replace("[^a-zA-Z#]", " ")
```

Out[11]:

	Date	User	Tweet	clean_tweet
0	2023-03-30 23:59:08+00:00	elonmusk	@mtaibbi 🚀	mtaibbi
1	2023-03-30 22:02:28+00:00	elonmusk	@iamharaldur !	iamharaldur
2	2023-03-30 07:44:08+00:00	elonmusk	@Galaxygamerone Coming soon	Galaxygamerone Coming soon
3	2023-03-30 07:15:48+00:00	elonmusk	@Tesmanian_com It will per vehicle	Tesmanian com It will per vehicle
4	2023-03-30 07:13:10+00:00	elonmusk	@TPostMillennial !!	TPostMillennial
5	2023-03-30 06:37:44+00:00	elonmusk	Old joke about agnostic technologists building...	Old joke about agnostic technologists building...
6	2023-03-29 23:18:40+00:00	elonmusk	@BillyM2k Makes me sad	BillyM k Makes me sad
7	2023-03-29 22:53:09+00:00	elonmusk	@GRDecter Combined with excess government spen...	GRDecter Combined with excess government spen...
8	2023-03-29 22:51:48+00:00	elonmusk	@GRDecter Serious issue. US policy has been to...	GRDecter Serious issue US policy has been to...
9	2023-03-29 22:41:45+00:00	elonmusk	@boringcompany Making progress!	boringcompany Making progress
10	2023-03-29 21:50:32+00:00	elonmusk	@andst7 🚀	andst
11	2023-03-29 21:43:57+00:00	elonmusk	@EvaFoxU 6 years ago	EvaFoxU years ago
12	2023-03-29 19:06:50+00:00	elonmusk	@iheartmindy This deserves close scrutiny 😊	iheartmindy This deserves close scrutiny
13	2023-03-29 19:05:37+00:00	elonmusk	@BillyM2k AI optimizing for the greatest under...	BillyM k AI optimizing for the greatest under...
14	2023-03-29 18:31:33+00:00	elonmusk	@Jason Yeah	Jason Yeah
15	2023-03-29 18:02:34+00:00	elonmusk	@TwitterSafety Given that there are ~500 milli...	TwitterSafety Given that there are milli...
16	2023-03-29 17:36:31+00:00	elonmusk	@TheChiefNerd Leading AGI developers will not ...	TheChiefNerd Leading AGI developers will not ...
17	2023-03-29 07:21:39+00:00	elonmusk	@TeslaDiva99 Tesla car temp is automatically k...	TeslaDiva Tesla car temp is automatically k...
18	2023-03-29 07:15:58+00:00	elonmusk	@george__mack I thought both SpaceX and Tesla ...	george mack I thought both SpaceX and Tesla ...
19	2023-03-28 19:56:26+00:00	elonmusk	@alx Sick burn lol	alx Sick burn lol

In [12]: # Remove short words

```
df['clean_tweet'] = df['clean_tweet'].apply(lambda x: " ".join([w for w in x.split() if len(w) > 2]))
df.head(20)
```

Out[12]:

	Date	User	Tweet	clean_tweet
0	2023-03-30 23:59:08+00:00	elonmusk	@mtaibbi 🎉	mtaibbi
1	2023-03-30 22:02:28+00:00	elonmusk	@iamharaldur !	iamharaldur
2	2023-03-30 07:44:08+00:00	elonmusk	@Galaxygamerone Coming soon	Galaxygamerone Coming soon
3	2023-03-30 07:15:48+00:00	elonmusk	@Tesmanian_com It will per vehicle	Tesmanian will vehicle
4	2023-03-30 07:13:10+00:00	elonmusk	@TPostMillennial !!	TPostMillennial
5	2023-03-30 06:37:44+00:00	elonmusk	Old joke about agnostic technologists building...	joke about agnostic technologists building art...
6	2023-03-29 23:18:40+00:00	elonmusk	@BillyM2k Makes me sad	BillyM Makes
7	2023-03-29 22:53:09+00:00	elonmusk	@GRDector Combined with excess government spend...	GRDector Combined with excess government spend...
8	2023-03-29 22:51:48+00:00	elonmusk	@GRDector Serious issue. US policy has been to...	GRDector Serious issue policy been heavy hande...
9	2023-03-29 22:41:45+00:00	elonmusk	@boringcompany Making progress!	boringcompany Making progress
10	2023-03-29 21:50:32+00:00	elonmusk	@andst7 🎉	andst
11	2023-03-29 21:43:57+00:00	elonmusk	@EvaFoxU 6 years ago	EvaFoxU years
12	2023-03-29 19:06:50+00:00	elonmusk	@iheartmindy This deserves close scrutiny 😊	iheartmindy This deserves close scrutiny
13	2023-03-29 19:05:37+00:00	elonmusk	@BillyM2k AI optimizing for the greatest under...	BillyM optimizing greatest understanding unive...
14	2023-03-29 18:31:33+00:00	elonmusk	@Jason Yeah	Jason Yeah
15	2023-03-29 18:02:34+00:00	elonmusk	@TwitterSafety Given that there are ~500 milli...	TwitterSafety Given that there million tweets ...
16	2023-03-29 17:36:31+00:00	elonmusk	@TheChiefNerd Leading AGI developers will not ...	TheChiefNerd Leading developers will heed this...
17	2023-03-29 07:21:39+00:00	elonmusk	@TeslaDiva99 Tesla car temp is automatically k...	TeslaDiva Tesla temp automatically kept within...
18	2023-03-29 07:15:58+00:00	elonmusk	@george_mack I thought both SpaceX and Tesla ...	george mack thought both SpaceX Tesla each les...
19	2023-03-28 19:56:26+00:00	elonmusk	@alx Sick burn lol	Sick burn

In [13]: # Individual words considered as tokens

```
tokenized_tweet = df['clean_tweet'].apply(lambda x: x.split())
tokenized_tweet.head(20)
```

Out[13]:

```
0 [mtaibbi]
1 [iamharaldur]
2 [Galaxygamerone, Coming, soon]
3 [Tesmanian, will, vehicle]
4 [TPostMillennial]
5 [joke, about, agnostic, technologists, buildin...]
6 [BillyM, Makes]
7 [GRDecter, Combined, with, excess, government,...]
8 [GRDecter, Serious, issue, policy, been, heavy...]
9 [boringcompany, Making, progress]
10 [andst]
11 [EvaFoxU, years]
12 [iheartmindy, This, deserves, close, scrutiny]
13 [BillyM, optimizing, greatest, understanding, ...]
14 [Jason, Yeah]
15 [TwitterSafety, Given, that, there, million, t...]
16 [TheChiefNerd, Leading, developers, will, heed...]
17 [TeslaDiva, Tesla, temp, automatically, kept, ...]
18 [george, mack, thought, both, SpaceX, Tesla, e...]
19 [Sick, burn]
Name: clean_tweet, dtype: object
```

In [14]: # stem the words

```
from nltk.stem.porter import PorterStemmer
stemmer = PorterStemmer()

tokenized_tweet = tokenized_tweet.apply(lambda sentence:[stemmer.stem(word) for word in sentence])
tokenized_tweet.head(20)
```

Out[14]:

```
0 [mtaibbi]
1 [iamharaldur]
2 [galaxygameron, come, soon]
3 [tesmanian, will, vehicl]
4 [tpostmillenni]
5 [joke, about, agnost, technologist, build, art...]
6 [billym, make]
7 [grdecter, combin, with, excess, govern, spend...]
8 [grdecter, seriou, issu, polici, been, heavi, ...]
9 [boringcompani, make, progress]
10 [andst]
11 [evafoxu, year]
12 [iheartmindi, thi, deserv, close, scrutini]
13 [billym, optim, greatest, understand, univers,...]
14 [jason, yeah]
15 [twittersafeti, given, that, there, million, t...]
16 [thechiefnerd, lead, develop, will, heed, thi,...]
17 [tesladiva, tesla, temp, automat, kept, within...]
18 [georg, mack, thought, both, spacex, tesla, ea...]
19 [sick, burn]
Name: clean_tweet, dtype: object
```

```
In [15]: # combine into single sentence
for i in range(len(tokenized_tweet)):
    tokenized_tweet[i] = " ".join(tokenized_tweet[i])

df['clean_tweet'] = tokenized_tweet
df.head()
```

Out[15]:

	Date	User	Tweet	clean_tweet
0	2023-03-30 23:59:08+00:00	elonmusk	@mtaibbi 🌐	mtaibbi
1	2023-03-30 22:02:28+00:00	elonmusk	@iamharaldur !	iamharaldur
2	2023-03-30 07:44:08+00:00	elonmusk	@Galaxygamerone Coming soon	galaxygameron come soon
3	2023-03-30 07:15:48+00:00	elonmusk	@Tesmanian_com It will per vehicle	tesmanian will vehicl
4	2023-03-30 07:13:10+00:00	elonmusk	@TPostMillennial !!	tpostmillenni

In [16]: !pip install wordcloud

Defaulting to user installation because normal site-packages is not writeable
Requirement already satisfied: wordcloud in c:\users\user\appdata\roaming\python\python39\site-packages (1.8.2.2)
Requirement already satisfied: matplotlib in c:\users\user\appdata\roaming\python\python39\site-packages (from wordcloud) (3.6.3)
Requirement already satisfied: pillow in c:\programdata\anaconda3\lib\site-packages (from wordcloud) (9.2.0)
Requirement already satisfied: numpy>=1.6.1 in c:\programdata\anaconda3\lib\site-packages (from wordcloud) (1.21.5)
Requirement already satisfied: fonttools>=4.22.0 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (4.25.0)
Requirement already satisfied: packaging>=20.0 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (21.3)
Requirement already satisfied: pyparsing>=2.2.1 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (3.0.9)
Requirement already satisfied: cycler>=0.10 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: python-dateutil>=2.7 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (2.8.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\user\appdata\roaming\python\python39\site-packages (from matplotlib->wordcloud) (1.0.7)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\programdata\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.4.2)
Requirement already satisfied: six>=1.5 in c:\programdata\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib->wordcloud) (1.16.0)

Exploratory Data Analysis


```
In [19]: # Frequent words visualization Positive
all_words = " ".join([sentence for sentence, label in zip(df['clean_tweet'], df['label']) if label == 1])
from wordcloud import WordCloud
wordcloud = WordCloud(width=800, height=500, random_state=42, max_font_size=100).generate(all_words)

# Plot the graph
plt.figure(figsize=(10, 7))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```

```

-----  

KeyError                                                 Traceback (most recent call last)  

C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\indexes\base.py in get  

    _loc(self, key, method, tolerance)  

       3628             try:  

-> 3629                 return self._engine.get_loc(casted_key)  

       3630             except KeyError as err:  

  

C:\ProgramData\Anaconda3\lib\site-packages\pandas\_libs\index.pyx in pandas._  

    libs.index.IndexEngine.get_loc()  

  

C:\ProgramData\Anaconda3\lib\site-packages\pandas\_libs\index.pyx in pandas._  

    libs.index.IndexEngine.get_loc()  

  

pandas\_libs\hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHas  

    hTable.get_item()  

  

pandas\_libs\hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHas  

    hTable.get_item()  

  

KeyError: 'label'

```

The above exception was the direct cause of the following exception:

```

KeyError                                                 Traceback (most recent call last)  

~\AppData\Local\Temp\ipykernel_9116\596955699.py in <module>  

    1 # Frequent words visualization Positive  

----> 2 all_words = " ".join([sentence for sentence, label in zip(df['clean_t  

weet'], df['label']) if label == 0])  

       3  

       4 from wordcloud import WordCloud  

       5 wordcloud = WordCloud(width=800, height=500, random_state=42, max_fon  

t_size=100).generate(all_words)  

  

C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\frame.py in __getitem__  

    _  

    (self, key)  

       3503             if self.columns.nlevels > 1:  

       3504                 return self._getitem_multilevel(key)  

-> 3505             indexer = self.columns.get_loc(key)  

       3506             if is_integer(indexer):  

       3507                 indexer = [indexer]  

  

C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\indexes\base.py in get  

    _loc(self, key, method, tolerance)  

       3629                 return self._engine.get_loc(casted_key)  

       3630             except KeyError as err:  

-> 3631                 raise KeyError(key) from err  

       3632             except TypeError:  

       3633                 # If we have a listlike key, _check_indexing_error wi  

11 raise  

  

KeyError: 'label'

```



```
In [23]: # Extract the hashtag
def hashtag_extract(tweet):
    hashtags = []
    # Loop words in tweet
    for word in tweet:
        ht = re.findall(r"#(\w+)", tweet)
        hashtags.append(ht)
    return hashtags
```

```
In [24]: def hashtag_extract(tweet):
    hashtags = []
    # Loop words in tweet
    for word in tweet.split():
        ht = re.findall(r"#(\w+)", word)
        hashtags.extend(ht)
    return hashtags
```

```
In [25]: # extract hashings from non-racist/ sexist tweets
ht_positive = hashtag_extract(df['clean_tweet'][df['label']==0])
ht_positive
```

```
-----
AttributeError                                     Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_9116\3475454840.py in <module>
      1 # extract hashings from non-racist/ sexist tweets
----> 2 ht_positive = hashtag_extract(df['clean_tweet'][df['label']==0])
      3 ht_positive

~\AppData\Local\Temp\ipykernel_9116\3404438414.py in hashtag_extract(tweet)
      2     hashtags = []
      3     # Loop words in tweet
----> 4     for word in tweet.split():
      5         ht = re.findall(r"#(\w+)", word)
      6         hashtags.extend(ht)

C:\ProgramData\Anaconda3\lib\site-packages\pandas\core\generic.py in __getatt
r__(self, name)
  553         ):
  554             return self[name]
-> 555         return object.__getattribute__(self, name)
  556
  557     def __setattr__(self, name: str, value) -> None:
```

AttributeError: 'Series' object has no attribute 'split'

```
In [26]: def hashtag_extract(tweets):
    hashtags = []
    # Loop over tweets
    for tweet in tweets:
        # Loop over words in tweet
        for word in tweet.split():
            ht = re.findall(r"#(\w+)", word)
            hashtags.extend(ht)
    return hashtags
print(hashtag_extract)

<function hashtag_extract at 0x00000236476C6DC0>
```

```
In [27]: # extract hashtags from non-racist/ sexist tweets
ht_positive = hashtag_extract(df['clean_tweet'][df['label']==0].tolist())
ht_positive
```

```
Out[27]: ['sobrav', 'freespeech']
```

```
In [28]: import pandas as pd
import numpy as np
import chart_studio.plotly as py
import cufflinks as cf
import seaborn as sns
import plotly.express as px
%matplotlib inline

# Make Plotly work in your Jupyter Notebook
from plotly.offline import download_plotlyjs, init_notebook_mode, plot, iplot
init_notebook_mode(connected=True)
# Use Plotly locally
cf.go_offline()
```

```
In [29]: df.columns
```

```
Out[29]: Index(['Date', 'User', 'Tweet', 'clean_tweet', 'label'], dtype='object')
```

```
In [ ]:
```