

```
In [ ]: %pip install nltk
        %pip install vaderSentiment
```

```
Requirement already satisfied: nltk in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (3.8.1)
Requirement already satisfied: click in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from nltk) (8.1.7)
Requirement already satisfied: joblib in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from nltk) (1.4.2)
Requirement already satisfied: regex<=2021.8.3 in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from nltk) (2024.5.15)
Requirement already satisfied: tqdm in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from nltk) (4.66.4)
Requirement already satisfied: colorama in c:\users\rowen\appdata\roaming\python\python311\site-packages (from click->nltk) (0.4.6)
Note: you may need to restart the kernel to use updated packages.
Requirement already satisfied: vaderSentiment in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (3.3.2)
Requirement already satisfied: requests in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from vaderSentiment) (2.32.2)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from requests->vaderSentiment) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from requests->vaderSentiment) (3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from requests->vaderSentiment) (2.2.1)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\rowen\appdata\local\programs\python\python311\lib\site-packages (from requests->vaderSentiment) (2024.2.2)
Note: you may need to restart the kernel to use updated packages.
```

```
In [ ]: import pandas as pd
import nltk
import re
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
from nltk.stem import WordNetLemmatizer
nltk.download('wordnet')
nltk.download('punkt')
nltk.download('stopwords')
```

```
[nltk_data] Downloading package wordnet to
[nltk_data] C:\Users\rowen\AppData\Roaming\nltk_data...
[nltk_data] Package wordnet is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data] C:\Users\rowen\AppData\Roaming\nltk_data...
[nltk_data] Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\rowen\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
```

```
Out[ ]: True
```

```
In [ ]: data = pd.read_csv('test.csv')
data = data.head(30)
print(data.shape)
```

```
(30, 2)
```

```
In [ ]: example = data['tweet']
print(example)
```

```

0 #studiolife #aislife #requires #passion #dedic...
1 @user #white #supremacists want everyone to s...
2 safe ways to heal your #acne!! #altwaystohe...
3 is the hp and the cursed child book up for res...
4 3rd #bihday to my amazing, hilarious #nephew...
5 choose to be :) #momtips
6 something inside me dies ð??|ð??¿â?` eyes nes...
7 #finished#tattoo#inked#ink#loveitâ?xi.ð #â?xi...
8 @user @user @user i will never understand why...
9 #delicious #food #lovelife #capetown mannaep...
10 1000dayswasted - narcosis infinite ep.. make m...
11 one of the world's greatest spoing events #l...
12 half way through the website now and #allgoing...
13 good food, good life , #enjoy and ð???ð???ð?...
14 i'll stand behind this #guncontrolplease #se...
15 i ate,i ate and i ate...ð???ð??? #jamaisasth...
16 @user got my @user limited edition rain or sh...
17 & #love & #hugs & #kisses too! how...
18 ð???ð???ð??? #girls #sun #fave @ london, uni...
19 thought factory: bbc neutrality on right wing ...
20 hey guys tommorow is the last day of my exams ...
21 @user @user @user #levyrroni #recuerdos mem...
22 my mind is like ð???ð???ð???%ð??? but my body l...
23 never been this down on myself in my entire li...
24 check twitterww - trends: "trending worldwide...
25 i thought i saw a mermaid!!! #ceegee #smcr ...
26 chick gets fucked hottest naked lady
27 happy bday lucyâ?`â?`ð??? xoxo #love #beautifu...
28 haroldfriday have a weekend filled with sunbe...
29 @user @user tried that! but nothing - will try...
Name: tweet, dtype: object

```

```
In [ ]: # Punctuation and symbol removal
```

```
removed = []

for text in example:
    # Remove punctuation
    text = re.sub(r'^\w\s', '', text)
    # Remove symbols
    text = re.sub(r'[\d\%&\^@~\!@Ä]', '', text)
    #Single spaces between words
    text = re.sub(r'\s+', ' ', text).strip()
    removed.append(text)

# Print cleaned texts
for text in removed:
    print(text)
```

studiolife aishlife requires passion dedication willpower to find newmaterials
user white supremacists want everyone to see the new birds movie and heres why
safe ways to heal your acne altwaystoheal healthy healing
is the hp and the cursed child book up for reservations already if yes where if no wh
en harrypotter pottermore favorite
3rd bihday to my amazing hilarious nephew eli ahmir uncle dave loves you and misses
choose to be momtips
something inside me dies eyes ness smokeyeyes tired lonely sof grunge
finishedtattooinkedinkloveit thanksaleeee
user user user i will never understand why my dad left me when i was so young deep in
thefeels
delicious food lovelife capetown mannaepicture resturant
1000dayswasted narcosis infinite ep make me aware grinding neuro bass lifestyle
one of the worlds greatest spoing events lemans24 teamaudi
half way through the website now and allgoingwell very
good food good life enjoy and this is called garlic bread iloveit
ill stand behind this guncontrolplease senselessshootings taketheguns comicrelief sti
llsad
i ate ate and i ate jamaisasthi fish curry prawn hilsa foodfestival foodies
user got my user limited edition rain or shine set today user user user user
amp love amp hugs amp kisses too how to keep your baby parenting healthcare
girls sun fave london united kingdom
thought factory bbc neutrality on right wing fascism politics media blm brexit trump
leadership gt3
hey guys tommorow is the last day of my exams im so happy yay
user user user levyrrooni recuerdos memories recuerdos friends life triunfodelamor
my mind is like but my body like sleepy stillallin
never been this down on myself in my entire life
check twitterww trends trending worldwide 1114 am bst1 oscarpistorius2 diplomalliseli
lerayakta3
i thought i saw a mermaid ceegee smcr inshot girls cute summer blur sun
chick gets fucked hottest naked lady
happy bday lucy xoxo love beautiful pizza instagood mileycyrus demilovato
haroldfriday have a weekend filled with sunbeams everyone healthy weekend
user user tried that but nothing will try again know you loved 2 but the 3rd light my
fave ampdeep rivers

```
In [ ]: #Number and duplicated characters removal
cleaned = []

for texts in removed:
    # Remove numbers
    texts = re.sub(r'\d+', '', texts)
    # Remove duplicate characters
    texts = re.sub(r'(\.)\1+', r'\1', texts)
    cleaned.append(texts)

# Print cleaned texts
for texts in cleaned:
    print(texts)
```

studiolife aislife requires pasion dedication wilpower to find newmaterials
user white supremacists want everyone to se the new birds movie and heres why
safe ways to heal your acne altwaystoheal healthy healing
is the hp and the cursed child bok up for reservations already if yes where if no whe
n harypoter potermore favorite
rd bihday to my amazing hilarious nephew eli ahmir uncle dave loves you and mises
chose to be momtips
something inside me dies eyes nes smokeyeyes tired lonely sof grunge
finishedtatoinkedinkloveit thanksale
user user user i wil never understand why my dad left me when i was so young dep inth
efels
delicious fod lovelife capetown manaepicture resturant
dayswasted narcosis infinite ep make me aware grinding neuro bas lifestyle
one of the worlds greatest spoing events lemans teamaudi
half way through the website now and algoingwel very
god fod god life enjoy and this is caled garlic bread iloveit
il stand behind this guncontrolplease senseleshotings taketheguns comicrelief stilsad
i ate ate and i ate jamaisasthi fish cury prawn hilsa fodfestival fodies
user got my user limited edition rain or shine set today user user user user
amp love amp hugs amp kisses to how to kep your baby parenting healthcare
girls sun fave london united kingdom
thought factory bc neutrality on right wing fascism politics media blm brexit trump l
eadership gt
hey guys tomorow is the last day of my exams im so hapy yay
user user user levyroni recuerdos memories recuerdos friends life triunfodelamor
my mind is like but my body like slepy stilalin
never ben this down on myself in my entire life
check twiterw trends trending worldwide am bst oscarpistorius diplomaliselilerayakta
i thought i saw a mermaid cege smcr inshot girls cute sumer blur sun
chick gets fucked hottest naked lady
hapy bday lucy xoxo love beautiful piza instagod mileycyrus demilovato
haroldfriday have a wekend filed with sunbeams everyone healthy wekend
user user tried that but nothing wil try again know you loved but the rd light my fav
e ampdep rivers

```
In [ ]: # Tokenization
tokenized = [word_tokenize(text) for text in cleaned]

for tokens in tokenized:
    print(tokens)
```

```

['studiolife', 'aislife', 'requires', 'pasion', 'dedication', 'wilpower', 'to', 'fin
d', 'newmaterials']
['user', 'white', 'supremacists', 'want', 'everyone', 'to', 'se', 'the', 'new', 'bird
s', 'movie', 'and', 'heres', 'why']
['safe', 'ways', 'to', 'heal', 'your', 'acne', 'altwaystoheal', 'healthy', 'healing']
['is', 'the', 'hp', 'and', 'the', 'cursed', 'child', 'bok', 'up', 'for', 'reservation
s', 'already', 'if', 'yes', 'where', 'if', 'no', 'when', 'harypoter', 'potermore', 'f
avorite']
['rd', 'bihday', 'to', 'my', 'amazing', 'hilarious', 'nephew', 'eli', 'ahmir', 'uncl
e', 'dave', 'loves', 'you', 'and', 'mises']
['chose', 'to', 'be', 'momtips']
['something', 'inside', 'me', 'dies', 'eyes', 'nes', 'smokeyeyes', 'tired', 'lonely',
'sof', 'grunge']
['finishedtatoinkedinkloveit', 'thanksale']
['user', 'user', 'user', 'i', 'wil', 'never', 'understand', 'why', 'my', 'dad', 'lef
t', 'me', 'when', 'i', 'was', 'so', 'young', 'dep', 'inthefels']
['delicious', 'fod', 'lovelife', 'capetown', 'manaepicure', 'resturant']
['dayswasted', 'narcosis', 'infinite', 'ep', 'make', 'me', 'aware', 'grinding', 'neur
o', 'bas', 'lifestyle']
['one', 'of', 'the', 'worlds', 'greatest', 'spoing', 'events', 'lemans', 'teamaudi']
['half', 'way', 'through', 'the', 'website', 'now', 'and', 'algoingwel', 'very']
['god', 'fod', 'god', 'life', 'enjoy', 'and', 'this', 'is', 'caled', 'garlic', 'brea
d', 'iloveit']
['il', 'stand', 'behind', 'this', 'guncontrolplease', 'senselesshotings', 'takethegun
s', 'comicrelief', 'stilsad']
['i', 'atei', 'ate', 'and', 'i', 'ate', 'jamaisasthi', 'fish', 'cury', 'prawn', 'hils
a', 'fodfestival', 'fodies']
['user', 'got', 'my', 'user', 'limited', 'edition', 'rain', 'or', 'shine', 'set', 'to
day', 'user', 'user', 'user', 'user']
['amp', 'love', 'amp', 'hugs', 'amp', 'kises', 'to', 'how', 'to', 'kep', 'your', 'bab
y', 'parenting', 'healthcare']
['girls', 'sun', 'fave', 'london', 'united', 'kingdom']
['thought', 'factory', 'bc', 'neutrality', 'on', 'right', 'wing', 'fascism', 'politic
s', 'media', 'blm', 'brexit', 'trump', 'leadership', 'gt']
['hey', 'guys', 'tomorrow', 'is', 'the', 'last', 'day', 'of', 'my', 'exams', 'im', 's
o', 'hapy', 'yay']
['user', 'user', 'user', 'levyroni', 'recuerdos', 'memories', 'recuerdos', 'friends',
'life', 'triunfodelamor']
['my', 'mind', 'is', 'like', 'but', 'my', 'body', 'like', 'slepy', 'stilalin']
['never', 'ben', 'this', 'down', 'on', 'myself', 'in', 'my', 'entire', 'life']
['check', 'twiterw', 'trends', 'trending', 'worldwide', 'am', 'bst', 'oscarpistoriu
s', 'diplomaliselilerayakta']
['i', 'thought', 'i', 'saw', 'a', 'mermaid', 'cege', 'smcr', 'inshot', 'girls', 'cut
e', 'sumer', 'blur', 'sun']
['chick', 'gets', 'fucked', 'hottest', 'naked', 'lady']
['hapy', 'bday', 'lucy', 'xoxo', 'love', 'beautiful', 'piza', 'instagod', 'mileycyru
s', 'demilovato']
['haroldfriday', 'have', 'a', 'wekend', 'filed', 'with', 'sunbeams', 'everyone', 'hea
lthy', 'wekend']
['user', 'user', 'tried', 'that', 'but', 'nothing', 'wil', 'try', 'again', 'know', 'y
ou', 'loved', 'but', 'the', 'rd', 'light', 'my', 'fave', 'ampdep', 'rivers']

```

```

In [ ]: #Removing stopwords
stop_words = set(stopwords.words('english'))

filtered = []

for tokens in tokenized:
    filtered_sentence = [word for word in tokens if word.lower() not in stop_words]
    filtered_string = ' '.join(filtered_sentence)

```

```

        filtered.append(filtered_string)

for i, filtered_string in enumerate(filtered, 1):
    print(filtered_string)

studiolife aislife requires pasion dedication wilpower find newmaterials
user white supremacists want everyone se new birds movie heres
safe ways heal acne altwaystoheal healthy healing
hp cursed child bok reservations already yes harypoter potermore favorite
rd bihday amazing hilarious nephew eli ahmir uncle dave loves mises
chose momtips
something inside dies eyes nes smokeyeyes tired lonely sof grunge
finishedtatoinkedinkloveit thanksale
user user user wil never understand dad left young dep inthefels
delicious fod lovelife capetown manaepicture resturant
dayswasted narcosis infinite ep make aware grinding neuro bas lifestyle
one worlds greatest spoing events lemans teamaudi
half way website algoingwel
god fod god life enjoy caled garlic bread iloveit
il stand behind guncontrolplease senseleshotings taketheguns comicrelief stilsad
atei ate ate jamaisasthi fish cury prawn hilsa fodfestival fodies
user got user limited edition rain shine set today user user user
amp love amp hugs amp kises kep baby parenting healthcare
girls sun fave london united kingdom
thought factory bc neutrality right wing fascism politics media blm brexit trump lead
ership gt
hey guys tomorow last day exams im hapy yay
user user user levyroni recuerdos memories recuerdos friends life triunfodelamor
mind like body like slepy stilalin
never ben entire life
check twiterw trends trending worldwide bst oscarpistorius diplomaliselilerayakta
thought saw mermaid cege smcr inshot girls cute sumer blur sun
chick gets fucked hottest naked lady
hapy bday lucy xoxo love beautiful piza instagod mileycyrus demilovato
haroldfriday wekend filed sunbeams everyone healthy wekend
user user tried nothing wil try know loved rd light fave ampdep rivers

```

```

In [ ]: #Lemmatization
        lemmatizer = WordNetLemmatizer()
        lemmatized = []

        for filtered_string in filtered:
            lemmatized_words = [lemmatizer.lemmatize(word.lower()) for word in word_tokenize(f
            lemmatized_sentence = ' '.join(lemmatized_words).capitalize()
            lemmatized.append(lemmatized_sentence)

        # Print all Lemmatized strings
        for lemmatized_string in lemmatized:
            print(lemmatized_string)

```

Studioli life requires passion dedication willpower find new materials
 User white supremacist want everyone see new bird movie here
 Safe way heal acne anyway to heal healthy healing
 Harry cursed child book reservation already yes Harry Potter more favorite
 Birthday amazing hilarious nephew Eli Amir uncle Dave love me
 Chose mom tips
 Something inside my eye not smoky eyes tired lonely soft grunge
 Finished tattoo ink love it thanksale
 User user user will never understand dad left young depressed in the fells
 Delicious food love life Cape Town man a picture restaurant
 Days wasted narcosis infinite episode make aware grinding neurobiology lifestyle
 One world greatest sporting event Le Mans team audi
 Half way website algorithm
 God food god life enjoy called garlic bread I love it
 I'll stand behind gun control please senseless shootings take the guns comic relief still sad
 Ate ate ate jamaisa sthi fish curry prawn hilsa food festival fodies
 User got user limited edition rain shine set today user user user user
 Amp love amp hug amp ki keep baby parenting healthcare
 Girl sun fave London United Kingdom
 Thought factory because neutrality right wing fascism politics medium Black Lives Matter Brexit Trump leadership
 Hey guy tomorrow last day exam I'm happy yay
 User user user levyróni recuerdos memory recuerdos friend life triunfo del amor
 Mind like body like sleepy still alive
 Never been entire life
 Check Twitter trend trending worldwide best Oscar Pistorius diplomatiselilerayakta
 Thought saw mermaid Cece Smalls in shot girl cute summer blur sun
 Chick get fucked hottest naked lady
 Happy birthday Lucy xoxo love beautiful pizza Instagram Miley Cyrus demilovato
 Harold Friday weekend filed sunbeam everyone healthy weekend
 User user tried nothing will try know loved road light fave amp depressed river

```

In [ ]: analyzer = SentimentIntensityAnalyzer()

compound = []

for strings in lemmatized:
    sentiment = analyzer.polarity_scores(strings)['compound']
    compound.append(sentiment)

# Create a new dataframe with cleaned tweets and their compound scores
result = pd.DataFrame({
    'processed_tweets': lemmatized,
    'sentiment_score': compound
})

print(result)
  
```

	processed_tweets	sentiment_score
0	Studiolife aislife requires pasion dedication ...	0.0000
1	User white supremacist want everyone se new bi...	0.2023
2	Safe way heal acne altwaystoheal healthy healing	0.6808
3	Hp cursed child bok reservation already yes ha...	0.6908
4	Rd bihday amazing hilarious nephew eli ahmir u...	0.8934
5	Chose momtips	0.0000
6	Something inside dy eye ne smokeyeyes tired lo...	-0.6597
7	Finishedtatoinkedinkloveit thanksale	0.0000
8	User user user wil never understand dad left y...	0.0000
9	Delicious fod lovelife capetown manaepicure re...	0.5719
10	Dayswasted narcosis infinite ep make aware gri...	0.0000
11	One world greatest spoing event lemans teamaudi	0.6369
12	Half way website algoingwel	0.0000
13	God fod god life enjoy caled garlic bread iloveit	0.7506
14	Il stand behind guncontrolplease senseleshotin...	0.0000
15	Atei ate ate jamaisasthi fish cury prawn hilsa...	0.0000
16	User got user limited edition rain shine set t...	-0.2263
17	Amp love amp hug amp ki kep baby parenting hea...	0.8074
18	Girl sun fave london united kingdom	0.6908
19	Thought factory bc neutrality right wing fasci...	0.2732
20	Hey guy tomorrow last day exam im hapy yay	0.5267
21	User user user levyroni recuerdos memory recue...	0.4939
22	Mind like body like slepy stilalin	0.6124
23	Never ben entire life	0.0000
24	Check twiterw trend trending worldwide bst osc...	0.0000
25	Thought saw mermaid cege smcr inshot girl cute...	0.4588
26	Chick get fucked hottest naked lady	-0.6597
27	Hapy bday lucy xoxo love beautiful piza instag...	0.9201
28	Haroldfriday weekend filed sunbeam everyone hea...	0.4019
29	User user tried nothing wil try know loved rd ...	0.7783

The model derives this score based on the processsed tweets. Some tweets have a positive sentiment score showing positive emotions like happiness and love. Meanwhile, the model detected tweets with negative scores as sad, depressing, and angry. The model detects those tweets that got a 0 sentiment score to be neutral, and based on the tweets that got this score, it seems like they are just statements that do not express particular emotions or feelings.