```
import pandas as pd
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer

data = pd.read_excel("D:\\Internship Progress\\Final all 3
datasets.xlsx")

text_column = 'snippet'

analyzer = SentimentIntensityAnalyzer()

def classify_sentiment(text):
    sentiment_score = analyzer.polarity_scores(text)
    compound_score = sentiment_score['compound']

    if compound_score >= 0.05:
        return 'Positive'
    elif compound_score <= -0.05:
        return 'Negative'
    else:
        return 'Neutral'

data['sentiment'] = data[text_column].apply(classify_sentiment)

data.to_excel("D:\Internship Progress\Final all 3
datasetSENTIMENTS.xlsx", index=False)

print(data)

count_positive = 0
count_negative = 0
count_neutral = 0

for sentiment in data['sentiment']:
    if sentiment == 'Positive':
        count_positive += 1
    elif sentiment == 'Negative':
        count_negative += 1
    elif sentiment == 'Neutral':
        count_neutral += 1

print(f"Number of Positive Sentiments: {count_positive}")
print(f"Number of Negative Sentiments: {count_negative}")
print(f"Number of Neutral Sentiments: {count_neutral}")

per1 = 0
per2 = 0
per3 = 0
total = len(data)
postitive_response = (count_positive / total) * 100
negative_response = (count_negative / total) * 100
```

```
neutral_response = (count_neutral / total) * 100

print("postitive_response", postitive_response)
print("negative_response", negative_response)
print("neutral_response", neutral_response)
```

```
                                                 snippet  \
0   Jul 31, 2023 — ... food & beverages, staff ser...
1   Aug 3, 2022 — has opened a new outlet at Depar...
2   Sep 25, 2023 — Bangalore airport also earns 70...
3   Dec 15, 2020 — The Quad retail and food and be...
4   BLR Airport. @BLRAirport. The official account...
..                                                ...
85  Let us join together to make Bangalore City a ...
86  But in Bangalore, you also need to look sidewa...
87  Day 5 of #IsraelPalestineWar | Israel retakes ...
88  ... food to the people to support them physica...
89  Aug 31, 2023 — Cheers from the International L...

                                            highlighs sentiment
0   ['food', 'beverages', 'Kempegowda Internationa...   Neutral
1                                          ['food']  Positive
2         ['Bangalore airport', 'Food and beverages']   Neutral
3   ['food and beverage', 'Kempegowda Internationa...   Neutral
4   ['BLR Airport', 'Kempegowda International Airp...   Neutral
..                                                ...       ...
85                         ['Bangalore', 'Bengaluru']  Positive
86               ['Bangalore', 'food', 'restaurants']   Neutral
87                                               NaN  Negative
88                                          ['food']  Negative
89   ['International', 'Bangalore airport', 'flight']  Positive

[90 rows x 3 columns]
Number of Positive Sentiments: 53
Number of Negative Sentiments: 6
Number of Neutral Sentiments: 31
postitive_response 58.88888888888889
negative_response 6.666666666666667
neutral_response 34.44444444444444
```

```
import pandas as pd
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer

data = pd.read_excel("D:\Internship Progress\
immigration_cleaned.xlsx")
text_column = 'snippet'

analyzer = SentimentIntensityAnalyzer()

def classify_sentiment(text):
```

```
    sentiment_score = analyzer.polarity_scores(text)
    compound_score = sentiment_score['compound']

    if compound_score >= 0.05:
        return 'Positive'
    elif compound_score <= -0.05:
        return 'Negative'
    else:
        return 'Neutral'

data['sentiment'] = data[text_column].apply(classify_sentiment)

data.to_excel("D:\\Internship Progress\\Final all 3
datasetSENTIMENTS2.xlsx", index=False)

print(data)

count_positive = 0
count_negative = 0
count_neutral = 0

for sentiment in data['sentiment']:
    if sentiment == 'Positive':
        count_positive += 1
    elif sentiment == 'Negative':
        count_negative += 1
    elif sentiment == 'Neutral':
        count_neutral += 1

print(f"Number of Positive Sentiments: {count_positive}")
print(f"Number of Negative Sentiments: {count_negative}")
print(f"Number of Neutral Sentiments: {count_neutral}")

per1 = 0
per2 = 0
per3 = 0
total = len(data)
postitive_response = (count_positive / total) * 100
negative_response = (count_negative / total) * 100
neutral_response = (count_neutral / total) * 100

print("postitive_response", postitive_response)
print("negative_response", negative_response)
print("neutral_response", neutral_response)
```

```
                                                snippet  \
0   Bravo Mumbai Immigration &Customs! They've rem...
1   - Application form - Cover letter with itinera...
2   3AM immigration queues at Bangalore Airport. I...
3   Myy son studying in the US landed at Delhi air...
```

```
4   Bangalore airport experience was stellar - tru...
..                                                ...
95  ... International Airport). More than that, it...
96  Parsis are Irani Migrants who landed & settled...
97  I booked my flight for Bengaluru but I asked t...
98  Jan 10, 2020 — Dear Mohit, We would like to in...
99  I'm seeking your help regarding my father's pa...

                                          highlighs sentiment
0                ['Immigration', 'forms', 'airport']  Negative
1   ['Application form', 'flight', 'entry', 'airpo...  Negative
2        ['immigration', 'Bangalore Airport', 'form']   Neutral
3             ['airport', 'Bangalore', 'immigration']  Negative
4                ['Bangalore airport', 'immigration']  Negative
..                                                ...       ...
95  ['International Airport', 'arrivals', 'passport']  Negative
96                                     ['Migrants']   Neutral
97                                  ['no problem']  Positive
98  ['you can carry upto 2 litres of alcohol purch...  Positive
99   ['passport', 'airport immigration', 'departure']  Positive

[100 rows x 3 columns]
Number of Positive Sentiments: 32
Number of Negative Sentiments: 47
Number of Neutral Sentiments: 21
postitive_response 32.0
negative_response 47.0
neutral_response 21.0
```

```python
import pandas as pd
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer

data = pd.read_excel("D:\Internship Progress\
Hospitality_cleaned.xlsx")
text_column = 'snippet'

analyzer = SentimentIntensityAnalyzer()

def classify_sentiment(text):
    sentiment_score = analyzer.polarity_scores(text)
    compound_score = sentiment_score['compound']

    if compound_score >= 0.05:
        return 'Positive'
    elif compound_score <= -0.05:
        return 'Negative'
    else:
        return 'Neutral'

data['sentiment'] = data[text_column].apply(classify_sentiment)
```

```
data.to_excel("D:\\Internship Progress\\Final all 3
datasetSENTIMENTS3.xlsx", index=False)

print(data)

count_positive = 0
count_negative = 0
count_neutral = 0

for sentiment in data['sentiment']:
    if sentiment == 'Positive':
        count_positive += 1
    elif sentiment == 'Negative':
        count_negative += 1
    elif sentiment == 'Neutral':
        count_neutral += 1

print(f"Number of Positive Sentiments: {count_positive}")
print(f"Number of Negative Sentiments: {count_negative}")
print(f"Number of Neutral Sentiments: {count_neutral}")

per1 = 0
per2 = 0
per3 = 0
total = len(data)
postitive_response = (count_positive / total) * 100
negative_response = (count_negative / total) * 100
neutral_response = (count_neutral / total) * 100

print("postitive_response", postitive_response)
print("negative_response", negative_response)
print("neutral_response", neutral_response)
```

```
                                              snippet  \
0   May 22, 2019 — Dive into the luxury and unmatc...
1   Taj Bangalore is a mere walk away from the Kem...
2   Sep 14, 2018 — We're offering a full range of ...
3   Red Key is a budget-friendly 2-star hotel in B...
4   Sep 14, 2018 — We're offering a full range of ...
..                                                ...
94  for making this tour happen The security was t...
95  Thank you for your wonderful hospitality and s...
96  Free event invite on Amex Platinum card at ITC...
97  YCO contestants getting ready for cook off at ...
98  for your wonderful hospitality in Germany.    ...

                                            highlighs sentiment
0   ['hospitality', 'Bangalore', 'Kempegowda Inter...  Negative
1   ['Bangalore', 'International Airport', 'hotel'...  Negative
```

```
2                          ['Hospitality', 'airport']   Neutral
3    ['hotel in Bangalore', 'Airport', 'hospitality']  Positive
4                          ['Hospitality', 'airport']   Neutral
..                                                ...       ...
94                                    ['hospitality']  Positive
95                ['hospitality', 'Bengaluru airport']  Positive
96                  ['hotels in Bengaluru', 'airport']  Positive
97                           ['Bangalore', 'airport']  Positive
98                       ['hospitality', 'Bangalore']  Positive

[99 rows x 3 columns]
Number of Positive Sentiments: 58
Number of Negative Sentiments: 14
Number of Neutral Sentiments: 27
postitive_response 58.58585858585859
negative_response 14.14141414141414
neutral_response 27.27272727272727
```

```python
import nltk
from sklearn.feature_extraction.text import TfidfVectorizer
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
import pandas as pd

a = pd.read_excel("D:\Internship Progress\Food and Beverages
cleaned.xlsx")
data = pd.DataFrame(a)

column_name = 'snippet'
corpus = data[column_name].tolist()

tokenized_data = [word_tokenize(text) for text in corpus]

stop_words = set(stopwords.words('english'))
filtered_data = [[word.lower() for word in text if word.isalnum() and
word.lower() not in stop_words] for text in tokenized_data]

preprocessed_data = [' '.join(text) for text in filtered_data]

vectorizer = TfidfVectorizer()
tfidf_matrix = vectorizer.fit_transform(preprocessed_data)

feature_names = vectorizer.get_feature_names_out()
tfidf_scores = tfidf_matrix.sum(axis=0).A1

word_tfidf_pairs = [(word, score) for word, score in
zip(feature_names, tfidf_scores)]

word_tfidf_pairs.sort(key=lambda x: x[1], reverse=True)

top_keywords = word_tfidf_pairs[:10]
```

```
for keyword, score in top_keywords:
    print(f"{keyword}: {score}")
```

```
airport: 5.89533371218242
bangalore: 5.394136358716413
food: 4.823487634552764
bengaluru: 2.7701025085776556
international: 2.630780233823315
india: 2.1741076509102384
2023: 2.053593533435689
image: 1.8536207437200516
beverages: 1.728407148835584
blr: 1.6604784776489478
```

```
pip install nltk
```

```
Requirement already satisfied: nltk in c:\users\91831\anaconda3\lib\
site-packages (3.7)
Requirement already satisfied: joblib in c:\users\91831\anaconda3\lib\
site-packages (from nltk) (1.1.0)
Requirement already satisfied: tqdm in c:\users\91831\anaconda3\lib\
site-packages (from nltk) (4.64.1)
Requirement already satisfied: regex>=2021.8.3 in c:\users\91831\
anaconda3\lib\site-packages (from nltk) (2022.7.9)
Requirement already satisfied: click in c:\users\91831\anaconda3\lib\
site-packages (from nltk) (8.0.4)
Requirement already satisfied: colorama in c:\users\91831\anaconda3\
lib\site-packages (from click->nltk) (0.4.6)
Note: you may need to restart the kernel to use updated packages.
```

```
nltk.download('punkt')
```

```
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

```
True
```

```
nltk.download('stopwords')
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Unzipping corpora\stopwords.zip.
```

```
True
```

```
pip install python-rake
```

```
Requirement already satisfied: python-rake in c:\users\91831\
anaconda3\lib\site-packages (1.5.0)
Note: you may need to restart the kernel to use updated packages.
```

```
import pandas as pd
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk import FreqDist

a = pd.read_excel("D:\\Internship Progress\\Food and Beverages
cleaned.xlsx")
data = pd.DataFrame(a)

column_name = 'snippet'
corpus = data[column_name]

nltk.download('stopwords')
nltk.download('punkt')

stop_words = set(stopwords.words('english'))

def extract_keywords(text):
    words = word_tokenize(text)
    words = [word.lower() for word in words if word.isalnum() and
word.lower() not in stop_words]
    freq_dist = FreqDist(words)
    keywords = [word for word, freq in freq_dist.most_common(5)]
    return keywords

for i, text in enumerate(corpus):
    keywords = extract_keywords(text)
    data['Keywords'] = data[column_name].apply(extract_keywords)
data
data.to_excel("D:\\Internship Progress\\Food and Beverages
cleaned.xlsx")
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

```
import pandas as pd
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk import FreqDist

a = pd.read_excel("D:\\Internship Progress\\Hospitality_cleaned.xlsx")
data = pd.DataFrame(a)

column_name = 'snippet'
```

```
corpus = data[column_name]

nltk.download('stopwords')
nltk.download('punkt')

stop_words = set(stopwords.words('english'))

def extract_keywords(text):
    words = word_tokenize(text)
    words = [word.lower() for word in words if word.isalnum() and
word.lower() not in stop_words]
    freq_dist = FreqDist(words)
    keywords = [word for word, freq in freq_dist.most_common(5)]
    return keywords

for i, text in enumerate(corpus):
    keywords = extract_keywords(text)
    data['Keywords'] = data[column_name].apply(extract_keywords)
data
data.to_excel("D:\\Internship Progress\\Hospitality_cleaned.xlsx")
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```

```
import pandas as pd
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from nltk import FreqDist

a = pd.read_excel("D:\\Internship Progress\\immigration_cleaned.xlsx")
data = pd.DataFrame(a)

column_name = 'snippet'
corpus = data[column_name]

nltk.download('stopwords')
nltk.download('punkt')

stop_words = set(stopwords.words('english'))

def extract_keywords(text):
    words = word_tokenize(text)
    words = [word.lower() for word in words if word.isalnum() and
word.lower() not in stop_words]
    freq_dist = FreqDist(words)
    keywords = [word for word, freq in freq_dist.most_common(5)]
```

```
    return keywords

for i, text in enumerate(corpus):
    keywords = extract_keywords(text)
    data['Keywords'] = data[column_name].apply(extract_keywords)
data
data.to_excel("D:\\Internship Progress\\immigration_cleaned.xlsx")
```

```
[nltk_data] Downloading package stopwords to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to
[nltk_data]     C:\Users\91831\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
```