

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

```
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.
True
```

```
import pandas as pd
import spacy
from textblob import TextBlob
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
import matplotlib.pyplot as plt
import seaborn as sns
from rake_nltk import Rake
import nltk
from collections import Counter
import re
```

```
df = pd.read_csv(r'BA_Reviews.csv')
```

```
df.shape
```

```
(1000, 3)
```

```
df.describe()
```

```

count    1000.000000
mean      499.500000
std       288.819436
min         0.000000
25%       249.750000
50%       499.500000
75%       749.250000
max       999.000000
```

```
df.isnull().sum()
```

```

Unnamed: 0      0
Verified_tags    0
reviews         0
dtype: int64
```


```
df['reviews'].iloc[100]
```



```
' This review is for LHR-SYD-LHR. BA015 and BA016. Business class had a 1-2-1 configuration on 777 giving space, privacy and easy aisle access. Seat reclines into comfortable flat bed. Good quality pillow and duvet. Privacy with sliding door. Good storage space around the seat and charging point. Flights on time - no delays. Gluten free meals were very poor in quality and presentation. Warm processed meat between two bland slices of GF bread apparently is a business class breakfast. On two sectors the main course (no choice offered) was bland fish which was served in a single small tray covered with foil - in business class? Condensation had formed as the dish was reheated (probably from chilled) so the offering was sitting in a weak watery liquid. One crew member appeared genuinely apologetic as they meekly served one meal. Do BA also serve the same GF offering in economy, economy plus and first on a long haul flight? My advice if GF - take your own food or dine in the lounge if you are '
```

```
# Deleting unwanted symbols and numeric data
```

```
df['processed_reviews']=df['reviews'].str.replace(r'[^a-zA-Z]', ' ', regex=True)
```

```
df.head()
```



Unnamed: 0			Verified_tags	reviews	processed_reviews	
0	0	Not Verified	Singapore to Heathrow. Business class on an...	Singapore to Heathrow Business class on an...		
1	1	Not Verified	I reported my damaged/ruined suitcase 5 wee...	I reported my damaged ruined suitcase wee...		
2	2	Not Verified	On March 1st, I flew from Berlin to São Pau...	On March st I flew from Berlin to S o Pau...		
3	3	Not Verified	The WORST customer experience! British Airway...	The WORST customer experience British Airway...		
4	4	Not Verified	Had to cancel my flight months in advance d...	Had to cancel my flight months in advance d...		

Next steps: [Generate code with df](#) [View recommended plots](#)


```
df['processed_reviews'].iloc[100]

' This review is for LHR SYD LHR BA and BA Business class had a configuration on giving space privacy and easy aisle access Seat reclines into comfortable flat bed Good quality pillow and duvet Privacy with sliding door Good storage space around the seat and charging point Flights on time no delays Gluten free meals were very poor in quality and presentation Warm processed meat between two bland slices of GF bread apparently is a business class breakfast On two sectors the main course no choice offered was bland fish which was served in a single small tray covered with foil in business class Condensation had formed as the dish was reheated probably from chilled so the offering was sitting in a weak watery liquid One crew member appeared genuinely apologetic as they meekly served one meal Do BA also serve the same GF offering in economy economy plus and first on a long haul flight? My advice if GF - take your own food or dine in the lounge if you are '

# Handling unwanted spaces

df['processed_reviews_2']=df['processed_reviews'].str.replace(r'[\s]+', ' ', regex = True)

df.head()
```



Unnamed: 0			Verified_tags	reviews	processed_reviews	processed_reviews_2	
0	0	Not Verified	Singapore to Heathrow. Business class on an...	Singapore to Heathrow Business class on an...	Singapore to Heathrow Business class on an...	Singapore to Heathrow Business class on an A...	
1	1	Not Verified	I reported my damaged/ruined suitcase 5 wee...	I reported my damaged ruined suitcase wee...	I reported my damaged ruined suitcase wee...	I reported my damaged ruined suitcase weeks a...	
2	2	Not Verified	On March 1st, I flew from Berlin to São Pau...	On March st I flew from Berlin to S o Pau...	On March st I flew from Berlin to S o Pau...	On March st I flew from Berlin to S o Paulo w...	
3	3	Not Verified	The WORST customer experience! British Airway...	The WORST customer experience British Airway...	The WORST customer experience British Airway...	The WORST customer experience British Airways...	
4	4	Not Verified	Had to cancel my flight months in advance d...	Had to cancel my flight months in advance d...	Had to cancel my flight months in advance d...	Had to cancel my flight months in advance due...	

Next steps: [Generate code with df](#) [View recommended plots](#)

```
df['reviews'].iloc[100]

' This review is for LHR-SYD-LHR. BA015 and BA016. Business class had a 1-2-1 configuration on 777 giving space, privacy and easy aisle access. Seat reclines into comfortable flat bed. Good quality pillow and duvet. Privacy with sliding door. Good storage space around the seat and charging point. Flights on time - no delays. Gluten free meals were very poor in quality and presentation. Warm processed meat between two bland slices of GF bread apparently is a business class breakfast. On two sectors the main course (no choice offered) was bland fish which was served in a single small tray covered with foil - in business class? Condensation had formed as the dish was reheated (probably from chilled) so the offering was sitting in a weak watery liquid. One crew member appeared genuinely apologetic as they meekly served one meal. Do BA also serve the same GF offering in economy, economy plus and first on a long haul flight? My advice if GF - take your own food or dine in the lounge if you are '

df.drop(columns = ['Unnamed: 0', 'Verified_tags','reviews','processed_reviews'], inplace = True)

df.head()
```



index	processed_reviews_2
0	Singapore to Heathrow Business class on an A is a Premium Economy Plus experience From the economy class sized toilets to the badly designed coffin shaped Business class pods the experience is a substandard business class experience Business class pod s storage is a small drawer at the FOOT of the pod The privacy screens are opaque plastic sheets so you are kept awake by your neighbour as their screen lights up your pod When the seat is made into its coffin shaped bed the cold hard plastic seat moulds are an unpleasant surprise for those that turn in their sleep You have a choice of having your neighbour step over your calf height footrest hopefully not brushing your feet with theirs or having to step over your neighbour s feet footrest With the footrest down heaven forbid you try to lever yourself into a more comfortable position as the slightest pressure causes the footrest to collapse into its floor level position If you are in the aisle seat which is open to the aisle kiss goodbye to any privacy as not only will you be on view to all passengers the crew will lean you to both serve your neighbour their food and talk to your neighbour over your privacy screen
1	I reported my damaged ruined suitcase weeks ago and I m still awaiting a reply I ll think twice about booking with BA again as the Customer Service is shocking to say the least I ve followed up with British Airways with at least emails per week requesting an update but with no joy so far
2	On March st I flew from Berlin to S o Paulo with BA but encountered several issues My first flight was delayed by several hours causing me to miss my connection at London Heathrow Upon arriving at the airport after midnight I was given a hotel voucher However the hotel designated by BA rudely denied me check in claiming it was fully booked I had to return to Heathrow at AM and found a different hotel paying for the taxi out of my own pocket As a result of these complications I arrived in S o Paulo more than hours late Despite my efforts to assert my rights and seek reimbursement for the taxi BA s customer service has been extremely poor Their phone lines are non functional emails are only answered automatically and the appropriate channels for such requests have not been responsive since the first week of March This experience has left me extremely disappointed especially considering the high standard of service I once associated with BA
3	The WORST customer experience British Airways BA is effectively robbing people They are not here to help and show no compassion My partner and I fell ill and had a doctor s note stating we couldn t fly BA said they d hold our flight amount as a voucher valid for a year Fair enough Caveat to this which was not stated to us you must call them and it is not usable on holiday packages However using this voucher was a nightmare We can t use it for both our tickets because the original was long haul and they refuse to refund half the amount BA s customer service was awful and rude effectively stealing from us This policy must stop We feel held hostage with no solution
4	Had to cancel my flight months in advance due to a change of schedule I was flying with American Airlines and British Airways Selected seats and when I attempted to get a refund on the selected seats AA refunded me with no issue when BA refused despite both companies having similar policies Avoid them fly with anyone else Customer service is a joke taking over weeks to respond to my email and giving me no further justification than what I was told on the phone

Show 25 per page

Next steps:

[Generate code with df](#)

[View recommended plots](#)

```
# NLP
nlp = spacy.load('en_core_web_sm')

# lemmatization
def lemmatization(text):
    doc = nlp(text)
    lemmalist = [word.lemma_ for word in doc]
    return ' '.join(lemmalist)

df['lemma_reviews']=df['processed_reviews_2'].apply(lemmatization)

df.head()
```



	processed_reviews_2	lemma_reviews
0	Singapore to Heathrow Business class on an A ...	Singapore to Heathrow Business class on an a...
1	I reported my damaged ruined suitcase weeks a...	I report my damage ruin suitcase week ago an...
2	On March st I flew from Berlin to S o Paulo w...	on March st I fly from Berlin to S o Paulo w...
3	The WORST customer experience British Airways...	the bad customer experience British Airways ...
4	Had to cancel my flight months in advance due...	have to cancel my flight month in advance du...

Next steps:

[Generate code with df](#)

[View recommended plots](#)

```
# Removing the stopwords
def remove_stopwords(text):
    doc = nlp(text)
    no_stopwords_list = [word.text for word in doc if not word.is_stop]
    return ' '.join(no_stopwords_list)

df['final_reviews']=df['lemma_reviews'].apply(remove_stopwords)

df.head()
```

	processed_reviews_2	lemma_reviews	final_reviews
0	Singapore to Heathrow Business class on an A...	Singapore to Heathrow Business class on an a...	Singapore Heathrow Business class Premium E...
1	I reported my damaged ruined suitcase weeks a...	I report my damage ruin suitcase week ago an...	report damage ruin suitcase week ago m awai...
2	On March st I flew from Berlin to S o Paulo w...	on March st I fly from Berlin to S o Paulo w...	March st fly Berlin S o Paulo BA encounter ...
3	The WORST customer experience British Airways...	the bad customer experience British Airways	bad customer experience British Airways BA ...

Next steps:

[Generate code with df](#)[View recommended plots](#)

df['final_reviews'].iloc[100]

```

review LHR SYD LHR BA BA Business class configuration space privacy easy aisle access Seat recline comfortable flat bed good q
uality pillow duvet Privacy slide door good storage space seat charge point Flights time delay Gluten free meal poor quality prese
ntation Warm process meat bland slice GF bread apparently business class breakfast sector main course choice offer bland fish serv
e single small tray cover foil business class Condensation form dish reheat probably chill offering sit weak watery liquid crew me
mber appear genuinely apologetic meekly serve meal BA serve gf offer economy economy plus long haul flight advice GF food dine lou
nge care BA hour travel companion good regular meal tell choice main course flight Crew impression job cover basic welcome engage
usual hospitallity industrv waffle service little substance deliver'

```

df.drop(columns = ['processed_reviews_2', 'lemma_reviews'], inplace = True)

df.head()

	final_reviews
0	Singapore Heathrow Business class Premium E...
1	report damage ruin suitcase week ago m awai...
2	March st fly Berlin S o Paulo BA encounter ...
3	bad customer experience British Airways BA ...
4	cancel flight month advance change schedule...

Next steps:

[Generate code with df](#)[View recommended plots](#)

```

# Initialize VADER sentiment analyzer
analyzer = SentimentIntensityAnalyzer()

```

```

# Function to get sentiment using VADER
def get_vader_sentiment(text):
    scores = analyzer.polarity_scores(text)
    return scores['compound']

```

```

# Apply sentiment analysis
df['sentiment'] = df['final_reviews'].apply(get_vader_sentiment)

```

```

# Classify sentiments
df['sentiment_class'] = df['sentiment'].apply(lambda x: 'positive' if x > 0 else ('negative' if x < 0 else 'neutral'))

```

```

# Calculate percentages
total_reviews = len(df)
positive_reviews = len(df[df['sentiment_class'] == 'positive'])
negative_reviews = len(df[df['sentiment_class'] == 'negative'])
neutral_reviews = len(df[df['sentiment_class'] == 'neutral'])

```

```

positive_percentage = (positive_reviews / total_reviews) * 100
negative_percentage = (negative_reviews / total_reviews) * 100
neutral_percentage = (neutral_reviews / total_reviews) * 100

```

```

print(positive_reviews)
print(negative_reviews)
print(neutral_reviews)

```

```

560
430
10

```

```

print(f"Positive Reviews: {positive_percentage:.2f}%")
print(f"Negative Reviews: {negative_percentage:.2f}%")

```

```

print(f'Negative Reviews: {negative_percentage:.2f}%')
print(f'Neutral Reviews: {neutral_percentage:.2f}%')

Positive Reviews: 56.00%
Negative Reviews: 43.00%
Neutral Reviews: 1.00%

# Create a DataFrame for the chart
sentiment_data = pd.DataFrame({
    'Sentiment': ['Positive', 'Negative', 'Neutral'],
    'Percentage': [positive_percentage, negative_percentage, neutral_percentage]
})

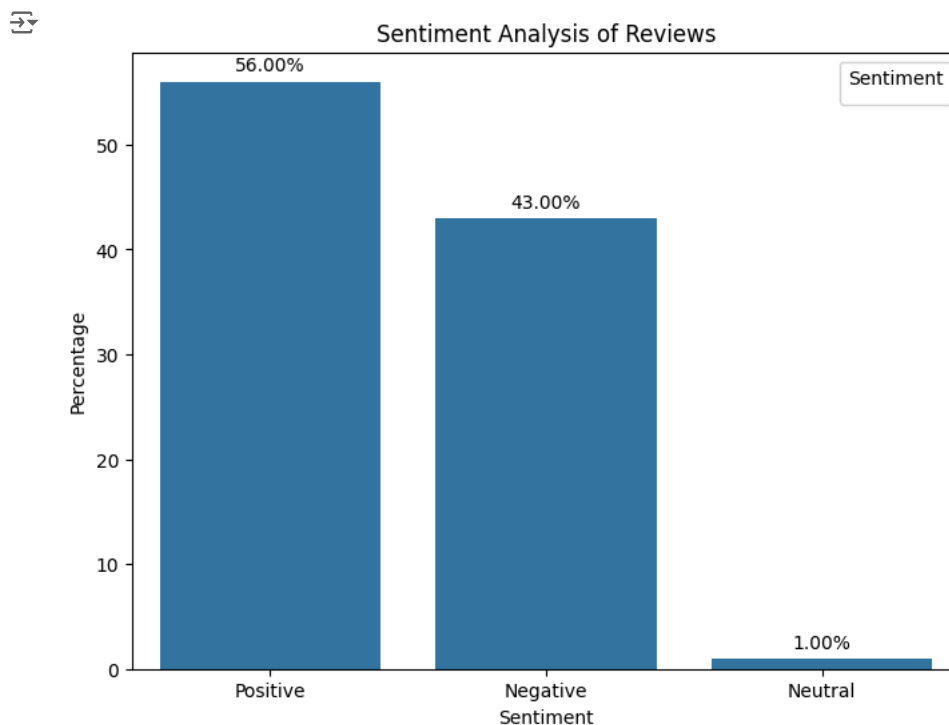
# Plot the chart
plt.figure(figsize=(8, 6))
barplot = sns.barplot(x='Sentiment', y='Percentage', data=sentiment_data)
plt.title('Sentiment Analysis of Reviews')
plt.ylabel('Percentage')
plt.xlabel('Sentiment')

# Add data labels
for index, row in sentiment_data.iterrows():
    barplot.text(row.name, row.Percentage + 1, f'{row.Percentage:.2f}%', color='black', ha="center")

# Add legend with colors
handles, labels = barplot.get_legend_handles_labels()
plt.legend(handles, labels, title='Sentiment', loc='upper right')

plt.show()

```



```

def preprocess_text(text):
    text = text.lower() # Convert to lowercase
    text = re.sub(r'^\w\s', '', text) # Remove punctuation
    return text

# Combine all reviews into a single string
all_reviews = ' '.join(df['final_reviews'].astype(str).apply(preprocess_text))

# Step 3: Count keywords
words = all_reviews.split()
keyword_counts = Counter(words)

```

```
# Step 4: Rank the keywords
ranked_keywords = keyword_counts.most_common()

# Display the top 10 keywords
for keyword, count in ranked_keywords[:10]:
    print(f'{keyword}: {count}')

flight: 2235
ba: 1329
seat: 883
service: 768
t: 707
time: 659
hour: 606

#a barchart showing keyword vs count

# Extract top 10 keywords and counts
top_keywords = [keyword for keyword, count in ranked_keywords[:10]]
top_counts = [count for keyword, count in ranked_keywords[:10]]

# Create bar chart
plt.figure(figsize=(10, 6))
plt.bar(top_keywords, top_counts)
plt.xlabel('Topics')
plt.ylabel('Frequency')
plt.title('Top 10 Topics')
plt.xticks(rotation=45, ha='right') # Rotate x-axis labels for readability
plt.tight_layout() # Adjust layout to prevent labels from overlapping
plt.show()
```

