

NLP Assignment - Problem Statement - 6

Group No - 26

Group Member Names:

1. Sunil Mittal (BITS ID : 2021SC04968)
2. Vikram Panwar (BITS ID : 2021SC04958)
3. Indira Saha (BITS ID : 2021sc04956)

```
In [92]: %%javascript
IPython.OutputArea.prototype._should_scroll = function(lines) {
    return false;
}
```

Importing All the required Libraries

```
In [93]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import re
import nltk
from nltk.corpus import stopwords
from nltk import word_tokenize, pos_tag
from nltk import Tree
from nltk.chunk import ne_chunk
from nltk.grammar import PCFG
from nltk.parse import ViterbiParser
#from nltk.grammar import toy_pcfg1 # toy_pcfg1 is one such parse tree created on dumm
#from nltk.grammar import toy_pcfg2
nltk.download('averaged_perceptron_tagger')
nltk.download('maxent_ne_chunker')
nltk.download('words')
```

```
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data] C:\Users\sunil\AppData\Roaming\nltk_data...
[nltk_data] Package averaged_perceptron_tagger is already up-to-
[nltk_data] date!
[nltk_data] Downloading package maxent_ne_chunker to
[nltk_data] C:\Users\sunil\AppData\Roaming\nltk_data...
[nltk_data] Package maxent_ne_chunker is already up-to-date!
[nltk_data] Downloading package words to
[nltk_data] C:\Users\sunil\AppData\Roaming\nltk_data...
[nltk_data] Package words is already up-to-date!
```

```
Out[93]: True
```

Read the CSV file

```
In [95]: # Read the CSV file into a DataFrame
file_path = 'chennai_reviews.csv'
df = pd.read_csv(file_path)
```

Describe the Data set head

```
In [96]: # Display the first few rows (head) of the DataFrame
df_head = df.head()
df_head
```

```
Out[96]:
```

	Hotel_name	Review_Title	Review_Text	Sentiment	Rating_Percentage	Unnamed: 5	Unnamed: 6
0	Accord Metropolitan	Excellent comfortableness during stay	Its really nice place to stay especially for b...	3	100	NaN	NaN
1	Accord Metropolitan	Not too comfortable	It seems that hotel does not check the basic a...	1	20	NaN	NaN
2	Accord Metropolitan	NaN	Worst hotel I have ever encountered. I will ne...	1	20	NaN	NaN
3	Accord Metropolitan	Best hotel	Had a good time in this hotel and the staff Ku...	3	100	NaN	NaN
4	Accord Metropolitan	NaN	good hotel and staff Veg food good non veg bre...	3	100	NaN	NaN

No. of rows, columns, and attributes. Plot the graph for the sentiment column

```
In [97]: # Get the number of rows and columns
num_rows, num_columns = df.shape
# Get the column names (attributes)
column_names = df.columns.tolist()
print ("No of Rows : " , num_rows)
print ("No of Cols : " , num_columns)
print ("List of Columns : " , column_names)
```

No of Rows : 4768
No of Cols : 9
List of Columns : ['Hotel_name', 'Review_Title', 'Review_Text', 'Sentiment', 'Rating_Percentage', 'Unnamed: 5', 'Unnamed: 6', 'Unnamed: 7', 'Unnamed: 8']

```
In [98]: df.describe()
```

Out[98]:

	Unnamed: 5	Unnamed: 7	Unnamed: 8
count	64.000000	1.0	1.0
mean	71.187500	2.0	60.0
std	21.138902	NaN	NaN
min	3.000000	2.0	60.0
25%	59.250000	2.0	60.0
50%	71.000000	2.0	60.0
75%	86.000000	2.0	60.0
max	100.000000	2.0	60.0

	Unnamed: 5	Unnamed: 7	Unnamed: 8
count	64.000000	1.0	1.0
mean	71.187500	2.0	60.0
std	21.138902	NaN	NaN
min	3.000000	2.0	60.0
25%	59.250000	2.0	60.0
50%	71.000000	2.0	60.0
75%	86.000000	2.0	60.0
max	100.000000	2.0	60.0

In [99]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4768 entries, 0 to 4767
Data columns (total 9 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Hotel_name            4768 non-null   object
1   Review_Title          4341 non-null   object
2   Review_Text           4767 non-null   object
3   Sentiment             4766 non-null   object
4   Rating_Percentage     4767 non-null   object
5   Unnamed: 5            64 non-null     float64
6   Unnamed: 6            3 non-null      object
7   Unnamed: 7            1 non-null      float64
8   Unnamed: 8            1 non-null      float64
dtypes: float64(3), object(6)
memory usage: 335.4+ KB
```

In [100...]

```
missing_val_df = pd.DataFrame().from_records([{'Column Name': col,
                                                'Missing Values': len(df[df[col].isna()]),
                                                'Missing Values (%)': np.round(len(df[df[col].isna()]) / len(df) * 100, 2)},
                                              for col in df.columns])

print('Missing values :')
missing_val_df
```

Missing values :

Out[100]:

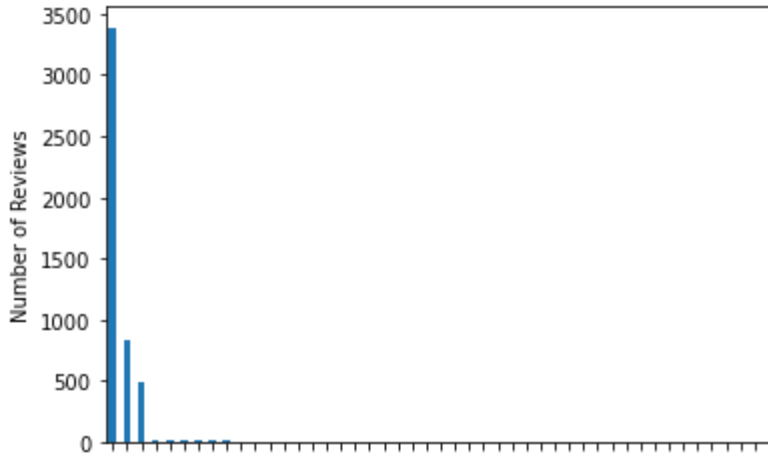
	Column Name	Missing Values	Missing Values (%)
0	Hotel_name	0	0.00
1	Review_Title	427	8.96
2	Review_Text	1	0.02
3	Sentiment	2	0.04
4	Rating_Percentage	1	0.02
5	Unnamed: 5	4704	98.66
6	Unnamed: 6	4765	99.94
7	Unnamed: 7	4767	99.98
8	Unnamed: 8	4767	99.98

Plotting the Sentiment Column

In [103...]

```
# Plot the sentiment column
sentiment_counts = df["Sentiment"].value_counts()
sentiment_counts.plot(kind='bar', title='Sentiment Distribution')
plt.xlabel('Sentiment')
plt.ylabel('Number of Reviews')
plt.show()
```

Sentiment Distribution



the service is mind blowing! Yummy and delicious breakfast! Complimentary cake for the wedding couples! A must place to try hop has flavours you haven't tasted before. More than food the rooms are tidy, bed clean and cosy, bathrooms modern with good water pressure polite and welcoming. Service was perfect. Food was tasty. Hotel has a small and beautiful garden. I would recommend this hotel for a short stay in-house but they do arrange for packed breakfast for you in the morning. The room are neat and clean and gives enough of peace in the night property is not very well managed. Most of the time only veg food was available and the concept of service apartment was not so greatly utilized. My good, the complimentary breakfast which impress all of us especially my father. Overall nice experience with calm and pleasant environment at home for his other uncompromising staff, who did not even care when we said we had a small wedding and wanted to change double the room I rent to the American consulate and major shopping area and malls. Bus stops, all and Autos available 24 hrs. Highly recommend to budget travellers and a room through goibibo and got a discount of 30%. They provided a complimentary breakfast, and the items in the restaurant were delicious. I Nagar so well in the centre of the city. They have also started wood oven fresh pizza which are too good. Room is clean though a little compact. Fine toast, cornflakes, eggs, fruit juice and coffee. It was halfway from anywhere. It was happy to come to think of it in conjunction with the price. My friends, I was travelling a long distance. I advice all not to miss this hotel. One thing I did not like from the hotel is that the places were untidy. I room service were handled promptly. Packed breakfast for an early morning night was really good. Recommended, especially for business travel.

AT AND CLEAN VERY GOOD CLANNES IN THER AND FOOD IS VERY TASTY AND GEORGIOUS AND IT HAS VERY GOOD SERVING THEM TO US VERY GOOD a wedding in the nearby vicinity. Good, smart and comfortable rooms with the price and particularly the breakfast was nice. Very courteous staff the room was very nice and clean. They provide hot water to bath all the things was very nice. There it was my very nice experience in this hotel and hygienic hotel. It was a nice experience in the hotel. The ambience was good and their staff was very polite and helpful. Value for money stay el provides room at cheap price. I like this hotel a lot. This hotel is good to stay with family and couple. Food facility of this hotel is also awesome. The rooms are good and well maintained. The service is good. But the reception office and booking has turned out to be very bad now a days. Very good. The quality and ambience of the hotel and the rooms were good and nice. Had a good time there. I had gone there for a holiday visit. soothing. We enjoyed very well. The overall experience of stay was fantastic. The exhibition center in front of this hotel was very beneficial for us. a smile. I am liked it. Food is very very nice. Sambaandly is awesome and portgal also extraordinary. I am feel happy and refreshing with that hotel. problem during check-in and check-out and there were no activities. The hotel room tariff was value for money and it is recommendable to others.

nd also a bar facility at night. The food is very good and especially chicken items were good. It is economical hotel and worth paying for the hotel. entire. A hub of Chennai! Breakfast tastes good. Food is a bit expensive. Can prefer eating in food court of Ascendas than going for a restaurant. read of breakfast. The staff were very warm, friendly and helpful. Will stay here again on my next trip to Chennai. A nice Hotel to stay in Chennai. they do not come inside

ge flights in Chennai. The hotel is close to the airport and provides free friendly pick up and drop. Friendly and helpful also in providing amenities. He to stay with family. Only for male that I go stay one or two days. If you want to stay here, book through makemytrip. You will get good discount. The room was being cleaned. We can visit in late well as really different to other makemytrip. I can get more than expected. Visiting again property. ea and location. We can visit in late well as really different to other makemytrip. I can get more than expected. Visiting again property. but it did not meet my expectations. It took them 45 minutes to allocate a room. A manager was very small and per 3 star standards is almost 20 km away from the city centre. But in the close proximity of 11 corridor I stayed there for 2 days and my overall experience was good itoes. AC is little bit order, freeze is very less. Wi-Fi is ok. Location is good middle of all H offices. Food and service is awesome. No words about it one of the best place in Chennai. Many people prefer this hotel only. All facilities are available in this hotel. Food is very good, in taste and healthy. alpful. Any problem with the room is attended promptly. The hotel however does not have a restaurant, but one can eat at many eateries close by is or not much use. I would advice if one want on business trip wants to stay for a day or two, it is nice to stay here as it is very near to Bus Stand.

Nice island tucked away in busy/noisy Chennai. Peace Park Inn is one of the best hotel. Their Chinese food is not bad either. The ice cream is honking or vehicles moving at fast pace which is the USP of this hotel. Enjoyed my stay here and it's easily accessible for my friends too to visit me when I am in Chennai. No food options available at a nice place on Mahabalipuram road in Chennai. However, the hotel did not had a any restaurant and you had to manage your food in the room only. The services were ok, however, the time and we asked for some extra grace time. Initially the staff was uncompromising on our situation. We find the hotel was centrally located just behind the central railway station and IN stadium. The service was provided stayed in this hotel. This is a good hotel with good facilities. There is a restaurant which offers tasty and varieties of food but a bit costly. The location is kind of good. The room is also good with TV, safe, air conditioning, toiletries, towels and tea set was available. The staff are very good. One of the very best location and very secured. Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

Very close to its always satisfactory to stay at this hotel. The staff is courteous and the service is good. The food is very good. The complimentary breakfast has a decent spread of south Indian dishes. Its in an extremely comfortable & clean hotel where I stayed was full of entertainment along with good cuisines. The infrastructure is very awesome and I enjoyed it a lot with An

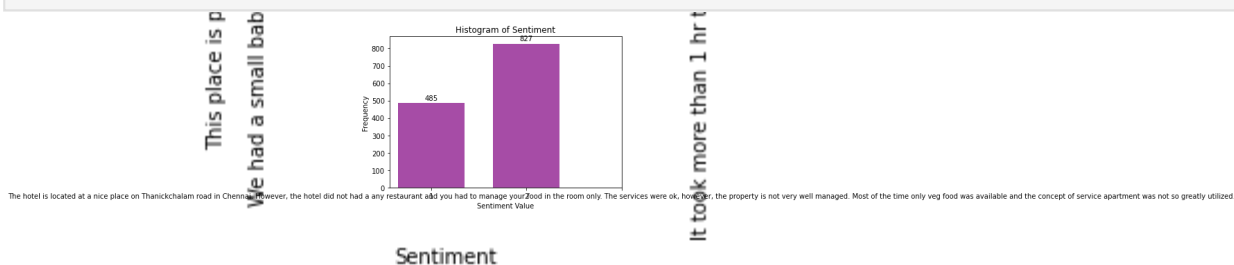
In [104...

```
# Plot histogram for the Sentiment column
# Calculate the histogram data
values, bins, patches = plt.hist(df['Sentiment'].dropna(), bins=np.arange(0.5, 3, 1),
#plt.figure(figsize=(20, 16))
# Annotate each bar with its height value
for i in range(len(values)):
    plt.text(bins[i] + 0.5, values[i] + 10, str(int(values[i])), ha='center', va='bottom')

plt.title("Histogram of Sentiment")
plt.xlabel("Sentiment Value")
plt.ylabel("Frequency")
plt.xticks([1, 2, 3])
plt.grid(False)
plt.show()

# Count rows by value in the Sentiment column
sentiment_counts = df['Sentiment'].value_counts().sort_index()

# Display the counts in a table format
sentiment_counts_table = pd.DataFrame(sentiment_counts).reset_index()
sentiment_counts_table.columns = ["Sentiment Value", "Count"]
sentiment_counts_table
```



Out[104]:

	Sentiment Value	Count
0	they do not come inside	1
1	as everything served gives a feeling of havin...	1
2	but it did not meet my expectations. It took ...	1
3	comfy rooms	1
4	excellent food. clean rooms. for some reason ...	1
5	food is very good.	1
6	in good location with all access to malls and...	4
7	the service is mind blowing! Yummy and delici...	4
8		1 485
9		2 827
10		3 3391
11	Akash Inn is one of the best place in Chennai....	1
12	An extremely comfortable & clean hotel with co...	1
13	I don't think it's a good match. Much costlier...	1
14	I had visited this hotel for my personal work ...	1
15	I visited Chennai recently and stayed in this ...	1
16	I visited the place with my friends in a group...	1
17	I was not satisfied with the service and time ...	1
18	If you make a plan for tour, then Chennai is o...	1
19	It is good place to stay in case you are trave...	1
20	It is located very convenient and room was spa...	1
21	It took more than 1 hr time to find out this h...	1
22	It was a nice hotel pleasant atmosphere, nice ...	1
23	It was the one of the best hotel of Chennai. T...	1
24	It's beside on OMR road, Thoraipakkam. Room se...	1
25	Its always satisfactory to stay at this hotel....	1
26	L SIGHT SEEING AND IN A COOL CLIMATE AND IT HA...	1
27	Location is very good. The staff are very good...	1
28	Nice hotel for a short visit specifically if o...	1
29	Nice interiors, good service and affordable st...	1
30	Nice island tucked away in busy/noisy Chennai....	4
31	Peace Park Inn is one of the best hotel. The r...	4
32	Stayed here for three nights. The hotel is wel...	1

	Sentiment Value	Count
33	The address is a very different and peaceful p...	1
34	The hotel is clean, the rooms are comfortable ...	1
35	The hotel is located at a nice place on Thanic...	4
36	The hotel is suitable for corporate guys for i...	1
37	The hotel where I stayed was full of entertain...	1
38	The room maintenance was very good they were c...	1
39	The room was small but TV, safe, air condition...	1
40	The room was very good. The room service was v...	1
41	This place is perfectly located in the heart o...	4
42	This was a one night stay to change flights in...	1
43	We had a small baby accompanying us for the vi...	1
44	We had gone to Chennai for getting our Son's a...	1
45	We have visited for Sankar Netralaya for my fa...	2
46	y enjoyed the one night stay at Gokulam Park S...	1

Remove all tags and special characters, stop words from the Dataset in the 'Review_Text' column

In [106...

```
# Initialize the list of stop words
stop_words = set(stopwords.words('english'))
# Function to clean the text
def clean_text(text):
    # Remove HTML tags using regular expression
    text = re.sub('<.*?>', '', text)
    # Remove special characters and numbers
    text = re.sub('[^a-zA-Z\s]', '', text)
    # Convert text to Lowercase
    text = text.lower()
    # Remove stop words
    text = ' '.join([word for word in text.split() if word not in stop_words])
    return text
# Drop rows where 'Review_Text' is NaN
df = df.dropna(subset=['Review_Text'])
# Apply the clean_text function to the 'Review_Text' column
df['Cleaned_Review_Text'] = df['Review_Text'].apply(clean_text)
# Display the first few rows to show the cleaned 'Review_Text'
df[['Review_Text', 'Cleaned_Review_Text']].head(10)
```

Out[106]:

	Review_Text	Cleaned_Review_Text
0	Its really nice place to stay especially for b...	really nice place stay especially business tou...
1	It seems that hotel does not check the basic a...	seems hotel check basic amenities room handing...
2	Worst hotel I have ever encountered. I will ne...	worst hotel ever encountered never think stay ...
3	Had a good time in this hotel and the staff Ku...	good time hotel staff kumaraishwarya house kee...
4	good hotel and staff Veg food good non veg bre...	good hotel staff veg food good non veg breakfa...
5	Great hotel. The staff if welcoming. Helpful. ...	great hotel staff welcoming helpful always
6	The hotel is mostly clean, but bathrooms need ...	hotel mostly clean bathrooms need care first b...
7	Awesome hotel with all needed amenities and lu...	awesome hotel needed amenities luxury friendly...
8	Very nice \nComfortable	nice comfortable
9	The staff of the hotel were polite. The brick...	staff hotel polite brick oven chefs extremely ...

Display the HMM POS tagging on the first 4 rows of 'Review_Text' and display a sparse tree for the same

In [110...

```
# Function to display a sparse tree (parse tree) for a sentence
def display_parse_tree(sentence):
    # Tokenize and POS Tag
    tokenized_sentence = word_tokenize(sentence)
    pos_tags = pos_tag(tokenized_sentence)

    # Named Entity Recognition (NER) to generate a sparse tree
    parse_tree = ne_chunk(pos_tags)

    return parse_tree

# Extract the first 4 rows from the 'Review_Text' column
df['Review_Text'] = df['Cleaned_Review_Text']
first_four_reviews = df['Cleaned_Review_Text'].head(4)

# Initialize an empty dictionary to hold the POS tags and parse trees
review_pos_tags = {}
review_parse_trees = {}
# Tokenize and POS tag each review
tagged_reviews = [pos_tag(word_tokenize(review)) for review in first_four_reviews]

# Print the tokens and their POS tags
for tags in tagged_reviews:
    print(tags)
    print("\n")

# Display a tree for each tagged review
for tags in tagged_reviews:
    tree = Tree('S', tags)
    tree.pretty_print()

# # Apply POS tagging and generate parse trees for each of the first 4 reviews
# for i, review in enumerate(first_four_reviews):
#     # Tokenize the words in the review
#     tokens = word_tokenize(review)
```

```
#      # Get the POS tags
#      pos_tags = pos_tag(tokens)

#      # Get the sparse tree (parse tree)
#      parse_tree = display_parse_tree(review)

#      # Save the POS tags and parse trees in the dictionary
#      review_pos_tags[f'Review_{i+1}'] = pos_tags
#      review_parse_trees[f'Review_{i+1}'] = parse_tree

# # Display the POS tags for the first 4 reviews
# review_pos_tags
```

[('really', 'RB'), ('nice', 'JJ'), ('place', 'NN'), ('stay', 'NN'), ('especially', 'RB'), ('business', 'NN'), ('tourist', 'NN'), ('purpose', 'NN')]

[('seems', 'VBZ'), ('hotel', 'NN'), ('check', 'NN'), ('basic', 'JJ'), ('amenities', 'NNS'), ('room', 'NN'), ('handing', 'VBG'), ('room', 'NN'), ('traveller', 'NN'), ('phone', 'NN'), ('room', 'NN'), ('working', 'VBG'), ('created', 'VBN'), ('problem', 'NN'), ('hectic', 'JJ'), ('day', 'NN'), ('travel', 'NN'), ('would', 'MD'), ('like', 'VB'), ('relax', 'JJ'), ('work', 'NN'), ('pace', 'NN'), ('sent', 'VBD'), ('technician', 'NN'), ('kept', 'NN'), ('looking', 'VBG'), ('fix', 'JJ'), ('hour', 'NN'), ('get', 'NN'), ('room', 'NN'), ('changed', 'VBD')]

[('worst', 'JJ'), ('hotel', 'NN'), ('ever', 'RB'), ('encountered', 'VBD'), ('never', 'RB'), ('think', 'VBP'), ('stay', 'JJ'), ('this', 'JJ'), ('hotel', 'NN'), ('future', 'NN')]

[('good', 'JJ'), ('time', 'NN'), ('hotel', 'NN'), ('staff', 'NN'), ('kumaraishwarya', 'NN'), ('house', 'NN'), ('keeping', 'VBG'), ('guy', 'JJ'), ('sure', 'JJ'), ('name', 'NN'), ('people', 'NNS'), ('make', 'VBP'), ('stay', 'JJ'), ('pleasant', 'JJ'), ('happy', 'JJ'), ('days', 'NNS'), ('sure', 'VBP'), ('back', 'RB'), ('place', 'NN'), ('stay', 'NN'), ('chennaigood', 'VBD'), ('job', 'NN'), ('guys', 'NNS')]

S
_____|_____
really/RB nice/JJ place/NN stay/NN especially/RB business/NN tourist/NN purpose/NN

S
_____|_____
_____|_____
seems/VBZ hotel/NN check/NN basic/JJ amenities/NNS room/NN handing/VBG room/NN traveller/NN phone/NN room/NN working/VBG created/VBN problem/NN hectic/JJ day/NN travel/NN would/MD like/VB relax/JJ work/NN pace/NN sent/VBD technician/JJ kept/NN looking/VBG fix/JJ hour/NN get/NN room/NN changed/VBD

S
_____|_____
worst/JJ hotel/NN ever/RB encountered/VBD never/RB think/VBP stay/JJ this/JJ hotel/NN future/NN

S
_____|_____
good/JJ time/NN hotel/NN staff/NN kumaraishwarya/ house/NN keeping/VBG guy/JJ sure/JJ name/NN people/NNS make/VBP stay/JJ pleasant/JJ happy/JJ days/NNS sure/VBP back/RB place/NN stay/NN chennaigood/VBD job/NN guys/NNS

NN

Parse the first 4 rows of 'v2' using Viterbi Parser ([Use toy_pcfg1 and toy_pcfg2 to get the probabilistic context-free grammars; use the PCFG

suitable for each sentence]

```
In [111...] sentences = df['Review_Text'].head(4).dropna().tolist()
sentences

Out[111]: ['really nice place stay especially business tourist purpose',
'seems hotel check basic amenities room handing room traveller phone room working cr
eated problem hectic day travel would like relax work pace sent technician kept looki
ng fix hour get room changed',
'worst hotel ever encountered never think stay thiis hotel future',
'good time hotel staff kumaraishwarya house keeping guy sure name people make stay p
leasant happy days sure back place stay chennaigood job guys']

In [112...] # # Define toy_pcfg1 and toy_pcfg2
toy_pcfg1_str = """
S -> NP VP [1.0]
NP -> Det N [0.45] | N [0.25] | 'I' [0.15] | NP PP [0.15]
VP -> V NP [0.35] | V ADJP [0.3] | V [0.2] | V NP PP [0.15]
PP -> P NP [0.7] | P [0.3]
ADJP -> Adj [0.7] | Adv Adj [0.3]
Det -> 'the' [0.2] | 'this' [0.1] | 'a' [0.2] | 'an' [0.15] | 'my' [0.15] | 'its'
N -> 'chennaigood' [0.025] | 'kumaraishwarya' [0.025] | 'technician' [0.025] | 'tr
V -> 'changed' [0.025] | 'looking' [0.025] | 'check' [0.025] | 'kept' [0.025] | 'w
P -> 'to' [0.2] | 'for' [0.1] | 'in' [0.1] | 'of' [0.1] | 'at' [0.1] | 'on' [0.1]
Adj -> 'worst' [0.1] | 'amenities' [0.05] | 'fix' [0.05] | 'problem' [0.05] | 'bas
Adv -> 'really' [0.04] | 'especially' [0.06] | 'very' [0.05] | 'never' [0.05] | 'f
"""

# Define PCFG toy_pcfg2
toy_pcfg2_str = """
S -> NP VP [1.0]
NP -> Det N [0.5] | N [0.3] | NP PP [0.2]
VP -> V NP [0.35] | V ADJP [0.3] | V [0.2] | V NP PP [0.15]
PP -> P NP [0.8] | P [0.2]
ADJP -> Adj [0.7] | Adv Adj [0.3]

Det -> 'the' [0.4] | 'this' [0.15] | 'its' [0.3] | 'thiis' [0.15]
N -> 'chennaigood' [0.01] | 'kumaraishwarya' [0.01] | 'technician' [0.0279] | 'tra
V -> 'changed' [0.025] | 'looking' [0.025] | 'check' [0.025] | 'kept' [0.025] | 'w
P -> 'with' [0.1] | 'for' [0.2] | 'in' [0.2] | 'of' [0.05] | 'at' [0.05] | 'on' [0
Adj -> 'worst' [0.1] | 'amenities' [0.05] | 'fix' [0.05] | 'problem' [0.05] | 'bas
Adv -> 'really' [0.04] | 'especially' [0.06] | 'very' [0.05] | 'never' [0.05] | 'f
"""

# Try creating the grammars again
toy_pcfg1 = nltk.PCFG.fromstring(toy_pcfg1_str)
toy_pcfg2 = nltk.PCFG.fromstring(toy_pcfg2_str)

toy_pcfg1, toy_pcfg2

toy_pcfg1 = nltk.PCFG.fromstring(toy_pcfg1_str)
toy_pcfg2 = nltk.PCFG.fromstring(toy_pcfg2_str)

# # Display the parse trees for the first 4 reviews
# review_parse_trees_pcfg
# Tokenize and parse the sentences
parsed_trees = []

# Extract the first four rows of the 'Review_Text' column
```

```

for sentence in sentences:
    tokens = nltk.word_tokenize(sentence)

    # Try parsing with toy_pcfg1
    parser1 = ViterbiParser(toy_pcfg1)
    trees1 = list(parser1.parse(tokens))

    # Try parsing with toy_pcfg2
    parser2 = ViterbiParser(toy_pcfg2)
    trees2 = list(parser2.parse(tokens))

    if len(trees1) > 0:
        parsed_trees.append(trees1[0])
    elif len(trees2) > 0:
        parsed_trees.append(trees2[0])
    else:
        parsed_trees.append(None)

parsed_trees

```

Out[112]: [None, None, None, None]

Above code Parse the first 4 rows of 'Review_Text' using Viterbi Parser , We used toy_pcfg1 and toy_pcfg2 to get the probabilistic context-free grammars. We manually looked at the sentences and tried to get as accurate grammer as possible.

There are multiple reasons as below why toy_pcfg1 and toy_pcfg2 wasn't able to parse tree .

1. Limited Coverage: Being a "toy" version means they are not comprehensive and only represent a small subset of a language. They might not be able to generate or parse a wide range of sentences beyond their defined rules.
2. Over-simplification: Toy PCFGs might overlook the complexities and intricacies of a natural language, leading to inaccuracies in modeling.
Not Suitable for Real-world Applications: Due to their simplicity, toy PCFGs are not suitable for real-world natural language processing tasks. They're primarily for educational purposes.
3. May Not Capture Ambiguities: Natural languages are filled with ambiguities, and a simplified PCFG might not be equipped to handle or represent these ambiguities.
4. Lack of Probabilistic Fidelity: The probabilities associated with the production rules in toy PCFGs might be arbitrarily assigned or overly simplified, not truly representing the probabilistic nature of real-world language constructs.