	data  Review Liked  Wow Loved this place. 1  Crust is not good. 0  Not tasty and the texture was just nasty. 0
	3 Stopped by during the late May bank holiday of 1 4 The selection on the menu was great and so wer 1 995 I think food should have flavor and texture an 0 996 Appetite instantly gone. 0
[7]: [7]:	997 Overall I was not impressed and would not go b 0 998 The whole experience was underwhelming, and I 0 999 Then, as if I hadn't wasted enough of my life 0 1000 rows × 2 columns  data.head()#Top 5 rows of the data set  Review Liked
	Wow Loved this place. 1 Crust is not good. 0 Not tasty and the texture was just nasty. 0 Stopped by during the late May bank holiday of 1 The selection on the menu was great and so wer 1
8]: 8]:	data tail()#Last 5 rows of data set  Review Liked  995 I think food should have flavor and texture an 0  996 Appetite instantly gone. 0  997 Overall I was not impressed and would not go b 0  998 The whole experience was underwhelming, and I 0
	Then, as if I hadn't wasted enough of my life 0  data.info()#information of data set like, data type, memory usage <class 'pandas.core.frame.dataframe'=""> RangeIndex: 1000 entries, 0 to 999  Data columns (total 2 columns):  # Column Non-Null Count Dtype </class>
	dtypes: int64(1), object(1) memory usage: 15.8+ KB  data.describe()#statistical information of data set  Liked  count 1000.00000  mean 0.50000  std 0.50025
	min 0.00000 25% 0.00000 50% 0.50000 75% 1.00000 max 1.00000
.1]: .1]: .2]: .2]:	Liked 0 dtype: int64  data.duplicated()
.3]:	<pre>3 False 4 False 4 False 5  995 False 996 False 997 False 998 False 999 False Length: 1000, dtype: bool  value_counts=data['Liked'].value_counts() print(value_counts)</pre>
4]:	Liked  1 500  0 500  Name: count, dtype: int64  pip install matplotlib  Requirement already satisfied: matplotlib in c:\users\admin\anaconda3\lib\site-packages (3.8.0)Note: you may need to restart the kernel to use updated packages.  Requirement already satisfied: contourpy>=1.0.1 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (1.2.0)  Requirement already satisfied: cycler>=0.10 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (0.11.0)  Requirement already satisfied: fonttools>=4.22.0 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (4.25.0)
5]:	Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (1.4.4) Requirement already satisfied: numpy<2,>=1.21 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (1.26.4) Requirement already satisfied: packaging>=20.0 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (23.1) Requirement already satisfied: pillow>=6.2.0 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (10.2.0) Requirement already satisfied: pyparsing>=2.3.1 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (30.9) Requirement already satisfied: python-dateutil>=2.7 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib) (2.8.2) Requirement already satisfied: six>=1.5 in c:\users\admin\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)  import matplotlib.pyplot as plt import seaborn as sns  value_counts.plot(kind= 'bar',color=['purple' , 'yellow']) plt.title("Sentiment value counts")
	plt.xlabel('Linked') plt.ylabel('Count') plt.xticks(ticks=[0,1],labels=['Positive','Negative'],rotation=0) plt.show()  Sentiment value counts  500 - 4
	y 300 - 200
	pip install wordcloud  Requirement already satisfied: wordcloud in c:\users\admin\anaconda3\lib\site-packages (1.9.3)  Requirement already satisfied: numpy>=1.6.1 in c:\users\admin\anaconda3\lib\site-packages (from wordcloud) (1.26.4)
	Requirement already satisfied: pillow in c:\users\admin\anaconda3\lib\site-packages (from wordcloud) (10.2.0) Requirement already satisfied: matplotlib in c:\users\admin\anaconda3\lib\site-packages (from wordcloud) (3.8.0) Requirement already satisfied: contourpy>=1.0.1 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.2.0) Requirement already satisfied: cycler>=0.10 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib->wordcloud) (0.11.0) Requirement already satisfied: fonttools>=4.22.0 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib->wordcloud) (4.25.0) Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.4.4) Requirement already satisfied: packaginy>=20.0 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib->wordcloud) (23.1) Requirement already satisfied: pyparsiny>=2.3.1 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib->wordcloud) (3.0.9) Requirement already satisfied: python-dateutil>=2.7 in c:\users\admin\anaconda3\lib\site-packages (from matplotlib->wordcloud) (2.8.2) Requirement already satisfied: six>=1.5 in c:\users\admin\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib->wordcloud) (1.16.0) Note: you may need to restart the kernel to use updated packages.
	<pre>from wordcloud import WordCloud  combined_text= " ".join(data['Review'])     wordcloud = WordCloud(width= 900, height =500 ,background_color='purple').generate(combined_text) plt.figure(figsize = (12,6)) plt.imshow(wordcloud, interpolation = 'bilinear') plt.axis('off') plt.title('Word Cloud of Reviews') plt.show()</pre> Word Cloud of Reviews
	went time  solved
	Vegas  friend US  cold  now 01/2/a sandwich Pho perfect  sandwich Pho perfect  table  worder ful autober service  bland  chicken  server  got  wonder ful autober service  bland  chicken  dish  server  got  wonder ful autober service  bland  chicken  dish  server  got  wonder ful autober service  bland  chicken  dish  server  got  wonder ful autober service  bland  chicken  go back  recommend  excellent  got  buffet  water  selection  chicken  go back  recommend  excellent  chicken  dish  selection  go back  recommend  excellent  chicken  dish  selection  go back  recommend  excellent  chicken  go back  recommend  excellent
	took say atmosphere probably side side side side side side side side
	<pre>plt.bar(target_word_counts.keys(),target_word_counts.values(),color = ['blue','yellow','pink']) plt.xlabel('Words') plt.ylabel('Frequency') plt.title('Frequency of specific words in Reviews') plt.show()</pre> Frequency of specific words in Reviews
	80 - 60 - 20 - 40 -
	20 - Tood place restaurant
3]:	<pre>#Text preprocessing  #convert a data set into lower case lowercased_text= data['Review'].str.lower()  print(lowercased_text)</pre>
	and tasty and the texture was just nasty.  stopped by during the late may bank holiday of  the selection on the menu was great and so wer  955 i think food should have flavor and texture an  966 appetite instantly gone.  97 overall i was not impressed and would not go b  98 the whole experience was underwhelming, and i  99 then, as if i hadn't wasted enough of my life  Name: Review, Length: 1000, dtype: object  # Tokenization import althe
5]: 6]:	<pre>import nltk nltk.download('punkt')  [nltk_data] Downloading package punkt to [nltk_data]</pre>
	print(data['Tokens'])  0
9]:	Then, , as, if, I, had, n't, wasted, enough,   Name: Tokens, Length: 1000, dtype: object     data.info()     <class 'pandas.core.frame.dataframe'="">     RangeIndex: 1000 entries, 0 to 999     Data columns (total 3 columns):</class>
80]:	Z Token food was terrible.  Review #NAME? The food was terrible. I would not recommend this place. I love this place. I won't be back.
32]:	
	print(data['Review'])  0
	Name: Review, Length: 1000, dtype: object  data['Review'].value_counts()
	OMG the food was delicioso There is nothing authentic about this place the spaghetti is nothing special whatsoever Then as if I hadn t wasted enough of my life there they poured salt in the wound by drawing out the time it took to bring the check Name: count, Length: 992, dtype: int64  #Removing the stop words like this ,is ,are,was import re import nltk nltk.download('stopwords')  [nltk_data] Downloading package stopwords to
35]: 36]: 37]:	<pre>[nltk_data]</pre>
	print(data['Tokens'])  [Wow, Loved, place] [Crust, good]  [Not, tasty, texture, nasty]  [Stopped, late, May, bank, holiday, Rick, Stev  [The, selection, menu, great, prices]  [I, think, food, flavor, texture, lacking]  [Appetite, instantly, gone]  [Overall, I, impressed, would, go, back]  [The, whole, experience, underwhelming, I, thi  [The, I, wasted, enough, life, poured, salt,
0]:  1]:  2]:	Name: Tokens, Length: 1000, dtype: object  #stemming #stemming is the process of reducing the word into root or base word form by removing suffix # example : "programming" stemmed is program  from nltk.stem.porter import PorterStemmer from nltk.tokenize import word_tokenize  p_stemmer = PorterStemmer()
4]:	data['p_stemmed'] = data['Review'].apply(lambda x: ' ' .join([p_stemmer.stem(word) for word in word_tokenize(x)]))  print(data['p_stemmed'])  wow love thi place crust is not good not tasti and the textur wa just nasti stop by dure the late may bank holiday off ric the select on the menu wa great and so were th  think food should have flavor and textur and appetit instantli gone  appetit instantli gone
15]:	997 overal i wa not impress and would not go back 998 the whole experi wa underwhelm and i think we 999 then as if i hadn t wast enough of my life the Name: p_stemmed, Length: 1000, dtype: object  import nltk nltk.download('wordnet')  [nltk_data] Downloading package wordnet to [nltk_data] C:\Users\ADMIN\AppData\Roaming\nltk_data [nltk_data] Package wordnet is already up-to-date!  True
	<pre>from nltk.stem import wordNetLemmatizer from nltk.tokenize import word_tokenize from nltk.corpus import wordnet  lemmatizer = WordNetLemmatizer()  data['Lemmatizer']= data['Review'].apply(lambda x: ' ' .join([lemmatizer.lemmatize(word) for word in word_tokenize(x)]))  print(data['Lemmatizer'])</pre>
	Wow Loved this place Crust is not good Not tasty and the texture wa just nasty Stopped by during the late May bank holiday of The selection on the menu wa great and so were  995 I think food should have flavor and texture an Appetite instantly gone 996 Appetite instantly gone 997 Overall I wa not impressed and would not go back 998 The whole experience wa underwhelming and I th 999 Then a if I hadn t wasted enough of my life th Name: Lemmatizer, Length: 1000, dtype: object
[0]: [1]:	<pre>import re data['No_Numbers'] = data['Review'].apply(lambda x: re.sub(r'\d+','' ,x))  print(data['No_Numbers'])  0</pre>
2]:	Appetite instantly gone 997 Overall I was not impressed and would not go b 998 The whole experience was underwhelming and I 999 Then as if I hadn t wasted enough of my life Name: No_Numbers, Length: 1000, dtype: object  #removing special characters like @ % # , * data['cleaned'] = data['Review'].apply(lambda x: re.sub(r'[^A-Za-z0-9\s]','' ,x))  print(data['cleaned'])  0
	Crust is not good  Not tasty and the texture was just nasty  Stopped by during the late May bank holiday of The selection on the menu was great and so wer  I think food should have flavor and texture an  Appetite instantly gone  Overall I was not impressed and would not go b  The whole experience was underwhelming and I  The whole experience was underwhelming and I  The whole experience was underwhelming and I  Same: cleaned, Length: 1000, dtype: object
5]:	#expanding method #don't eat food in this hotel, when we apply expanded text it will convert into donot eat food in this hotel  !pip install contractions  Requirement already satisfied: contractions in c:\users\admin\anaconda3\lib\site-packages (0.1.73)  Requirement already satisfied: textsearch>=0.0.21 in c:\users\admin\anaconda3\lib\site-packages (from contractions) (0.0.24)  Requirement already satisfied: anyascii in c:\users\admin\anaconda3\lib\site-packages (from textsearch>=0.0.21->contractions) (0.3.2)  Requirement already satisfied: pyahocorasick in c:\users\admin\anaconda3\lib\site-packages (from textsearch>=0.0.21->contractions) (2.1.0)  import sys  Lisys executable\-m nin install contractions
57]:	<pre>!{sys.executable}-m pip install contractions  'C:\Users\ADMIN\anaconda3\python.exe-m' is not recognized as an internal or external command, operable program or batch file.  import contractions data['Expanded'] = data['Review'].apply(contractions.fix)  print(data['Expanded'])  0</pre>
	Stopped by during the late May bank holiday of The selection on the menu was great and so wer  1  95
60]: 61]:	Wow Loved this place Crust is not good Not tasty and the texture was just nasty
2]:	Stopped by during the late May bank holiday of  The selection on the menu was great and so wer  """""""""""""""""""""""""""""""""
54]: 55]:	Requirement already satisfied: beautifulsoup4 in c:\users\admin\anaconda3\lib\site-packages (4.12.2) Requirement already satisfied: soupsieve>1.2 in c:\users\admin\anaconda3\lib\site-packages (from beautifulsoup4) (2.5)  from bs4 import BeautifulSoup
	<pre>print(data['cleaned'])  0</pre>
8]:	Overall I was not impressed and would not go b  The whole experience was underwhelming and I  Provided the sperience was underwhelming and I  The as if I hadn t wasted enough of my life  Name: cleaned, Length: 1000, dtype: object   X = vectorizer.fit_transform(data['Review'])  print(X.toarray())  [[0. 0. 0 0. 0. 0.]  [[0. 0. 0 0. 0. 0.]  [[0. 0. 0 0. 0. 0.]  [[0. 0. 0 0. 0. 0.]
	<pre>[0. 0. 0 0. 0. 0.] [0. 0. 0 0. 0. 0.] [0. 0. 0 0. 0. 0.]  #building a machine learning model from sklearn.model_selection import train_test_split from sklearn.naive_bayes import MultinomialNB from sklearn.metrics import accuracy_score, classification_report  vectorizer = TfidfVectorizer()  X = vectorizer.fit_transform(data['Review'])</pre>
′0]:	<pre>X = vectorizer.fit_transform(data['Review']) Y = data['tiked']  X_train,X_test,Y_train,Y_test = train_test_split(X,Y , test_size = 0.2 ,random_state = 42)  model = MultinomialNB() model.fit(X_train,Y_train)  v MultinomialNB MultinomialNB()</pre>
70]: 71]: 72]:	
<pre>70]: 71]: 72]: 73]: 74]: 76]: 77]:</pre>	<pre>accuracy = accuracy_score(Y_test,Y_pred)</pre>

