In [1]:
```
# import the  data set and do the data preprocessing
```

In [ ]:
```
# problem statement
#Develop a asentiment analysis model in classify reviews as positive or negati
#Preprocess the review text using techniques such as lower casing, removing st
#use the trained model accurately predict the sentiment of new, unseen reviews
```

In [ ]:
```
import pandas as pd
```

In [2]:
```
data=pd.read_csv('Reviews.csv')
```

In [3]:
```
print(data)
```

```
                                                Review  Liked
0                            Wow... Loved this place.      1
1                                  Crust is not good.      0
2           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
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 x 2 columns]
```

In [4]:
```
data
```

Out[4]:

	Review	Liked
0	Wow... Loved this place.	1
1	Crust is not good.	0
2	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
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

In [5]:
```
#we can review the top 5 rows
data.head()
```

Out[5]:

	Review	Liked
0	Wow... Loved this place.	1
1	Crust is not good.	0
2	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

In [6]:
```
# top 10 rows
data.head(10)
```

Out[6]:

	Review	Liked
0	Wow... Loved this place.	1
1	Crust is not good.	0
2	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
5	Now I am getting angry and I want my damn pho.	0
6	Honeslty it didn't taste THAT fresh.)	0
7	The potatoes were like rubber and you could te...	0
8	The fries were great too.	1
9	A great touch.	1

In [7]:
```
data.tail() # last 5 rows
```

Out[7]:

	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
999	Then, as if I hadn't wasted enough of my life ...	0

In [8]:
```
# last 10 rows
data.tail(10)
```

Out[8]:

	Review	Liked
990	The refried beans that came with my meal were ...	0
991	Spend your money and time some place else.	0
992	A lady at the table next to us found a live gr...	0
993	the presentation of the food was awful.	0
994	I can't tell you how disappointed I was.	0
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
999	Then, as if I hadn't wasted enough of my life ...	0

In [9]:
```
data.info() # data set info like data types,entries of data set, memory usage.
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype 
---  ------  --------------  ----- 
 0   Review  1000 non-null   object
 1   Liked   1000 non-null   int64 
dtypes: int64(1), object(1)
memory usage: 15.8+ KB
```

In [10]:
```
data.isnull().sum() # checking the null values.
```

Out[10]:
```
Review    0
Liked     0
dtype: int64
```

In [11]:
```
data.duplicated() # checking the duplicated values.
```

Out[11]:
```
0      False
1      False
2      False
3      False
4      False
       ...
995    False
996    False
997    False
998    False
999    False
Length: 1000, dtype: bool
```

In [13]:
```
import seaborn as sns
import matplotlib.pyplot as plt
```

In [48]:
```
pip install wordcloud
```

```
Collecting wordcloudNote: you may need to restart the kernel to use updated
packages.

  Obtaining dependency information for wordcloud from https://files.pythonh
osted.org/packages/f5/b0/247159f61c5d5d6647171bef84430b7efad4db504f02296740
24f3a4f7f2/wordcloud-1.9.3-cp311-cp311-win_amd64.whl.metadata (https://file
s.pythonhosted.org/packages/f5/b0/247159f61c5d5d6647171bef84430b7efad4db504
f0229674024f3a4f7f2/wordcloud-1.9.3-cp311-cp311-win_amd64.whl.metadata)
  Downloading wordcloud-1.9.3-cp311-cp311-win_amd64.whl.metadata (3.5 kB)
Requirement already satisfied: numpy>=1.6.1 in c:\users\dk2100tx\anaconda3
\lib\site-packages (from wordcloud) (1.24.3)
Requirement already satisfied: pillow in c:\users\dk2100tx\anaconda3\lib\si
te-packages (from wordcloud) (9.4.0)
Requirement already satisfied: matplotlib in c:\users\dk2100tx\anaconda3\li
b\site-packages (from wordcloud) (3.7.1)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\dk2100tx\anacon
da3\lib\site-packages (from matplotlib->wordcloud) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\dk2100tx\anaconda3
\lib\site-packages (from matplotlib->wordcloud) (0.11.0)
```

In [49]:
```
from wordcloud import WordCloud
```

In [50]:
```
combined_text=" ".join(data['Review'])
```

In [53]:
```
wordcloud=WordCloud(width=800,height=400,background_color='white').generate(co
```

In [55]:
```
#plot the word cloud
plt.figure(figsize=(10,6))
plt.imshow(wordcloud,interpolation='bilinear')
plt.axis('off')
plt.title('Word Cloud of Reviews')
plt.show()
```

In [23]:
```
from collections import Counter
```

In [44]:
```
targeted_words = ['good','great','amazing','bad']
all_words=" ".join(data['Review']).lower().split()  #flatten reviews into a si
word_counts = Counter(all_words)
target_word_count = {word: word_counts[word] for word in targeted_words}
```

In [45]:
```
#plottting
plt.figure(figsize=(8,6))
plt.bar(target_word_count.keys(), target_word_count.values(), color=['blue','g
plt.xlabel('Words')
plt.ylabel('Frequency')
plt.title('Frequency of specific words in reviews')
plt.show()
```