Q4:Analyze and visualize sentiment patterns in social media data to understand public opinion and attitudes towards specific topics or brands.

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
df=pd.read_csv("twitter_training.csv")
df
```

Out[2]:

	2401	Borderlands	Positive	im getting on borderlands and i will murder you all ,
0	2401	Borderlands	Positive	I am coming to the borders and I will kill you...
1	2401	Borderlands	Positive	im getting on borderlands and i will kill you ...
2	2401	Borderlands	Positive	im coming on borderlands and i will murder you...
3	2401	Borderlands	Positive	im getting on borderlands 2 and i will murder ...
4	2401	Borderlands	Positive	im getting into borderlands and i can murder y...
...	...	...	...	...
74676	9200	Nvidia	Positive	Just realized that the Windows partition of my...
74677	9200	Nvidia	Positive	Just realized that my Mac window partition is ...
74678	9200	Nvidia	Positive	Just realized the windows partition of my Mac ...
74679	9200	Nvidia	Positive	Just realized between the windows partition of...
74680	9200	Nvidia	Positive	Just like the windows partition of my Mac is l...

74681 rows × 4 columns

In [3]:
```
df.head()
```

Out[3]:

	2401	Borderlands	Positive	im getting on borderlands and i will murder you all ,
0	2401	Borderlands	Positive	I am coming to the borders and I will kill you...
1	2401	Borderlands	Positive	im getting on borderlands and i will kill you ...
2	2401	Borderlands	Positive	im coming on borderlands and i will murder you...
3	2401	Borderlands	Positive	im getting on borderlands 2 and i will murder ...
4	2401	Borderlands	Positive	im getting into borderlands and i can murder y...

In [6]:
```
col=['ID', 'Entity', 'Sentiment', 'Content']
df=pd.read_csv("twitter_training.csv",names=col)
df.head()
```

Out[6]:

	ID	Entity	Sentiment	Content
0	2401	Borderlands	Positive	im getting on borderlands and i will murder yo...
1	2401	Borderlands	Positive	I am coming to the borders and I will kill you...
2	2401	Borderlands	Positive	im getting on borderlands and i will kill you ...
3	2401	Borderlands	Positive	im coming on borderlands and i will murder you...
4	2401	Borderlands	Positive	im getting on borderlands 2 and i will murder ...

In [7]:
```
df.shape
```

Out[7]:
```
(74682, 4)
```

In [8]:
```
df.ndim
```

Out[8]:
```
2
```

In [9]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 74682 entries, 0 to 74681
Data columns (total 4 columns):
 #   Column     Non-Null Count  Dtype 
---  ------     --------------  ----- 
 0   ID         74682 non-null  int64 
 1   Entity     74682 non-null  object
 2   Sentiment  74682 non-null  object
 3   Content    73996 non-null  object
dtypes: int64(1), object(3)
memory usage: 2.3+ MB
```

In [11]:
```
df.describe
```

Out[11]:
```
<bound method NDFrame.describe of          ID       Entity Sentiment  \
0      2401  Borderlands  Positive
1      2401  Borderlands  Positive
2      2401  Borderlands  Positive
3      2401  Borderlands  Positive
4      2401  Borderlands  Positive
...     ...          ...       ...
74677  9200       Nvidia  Positive
74678  9200       Nvidia  Positive
74679  9200       Nvidia  Positive
74680  9200       Nvidia  Positive
74681  9200       Nvidia  Positive

                                                 Content
0      im getting on borderlands and i will murder yo...
1      I am coming to the borders and I will kill you...
2      im getting on borderlands and i will kill you ...
3      im coming on borderlands and i will murder you...
4      im getting on borderlands 2 and i will murder ...
...                                                  ...
74677  Just realized that the Windows partition of my...
74678  Just realized that my Mac window partition is ...
74679  Just realized the windows partition of my Mac ...
74680  Just realized between the windows partition of...
74681  Just like the windows partition of my Mac is l...

[74682 rows x 4 columns]>
```

In [12]:
```
df.isnull().sum()
```

Out[12]:
```
ID             0
Entity         0
Sentiment      0
Content      686
dtype: int64
```

In [13]:
```
df.dropna(axis=0,inplace=True)
```

In [14]:
```
df.isnull().sum()
```

Out[14]:
```
ID           0
Entity       0
Sentiment    0
Content      0
dtype: int64
```

In [15]:
```
df.duplicated().sum()
```

Out[15]:
```
2341
```

In [16]:
```
df.drop_duplicates(inplace=True)
df.duplicated().sum()
```

Out[16]:
```
0
```

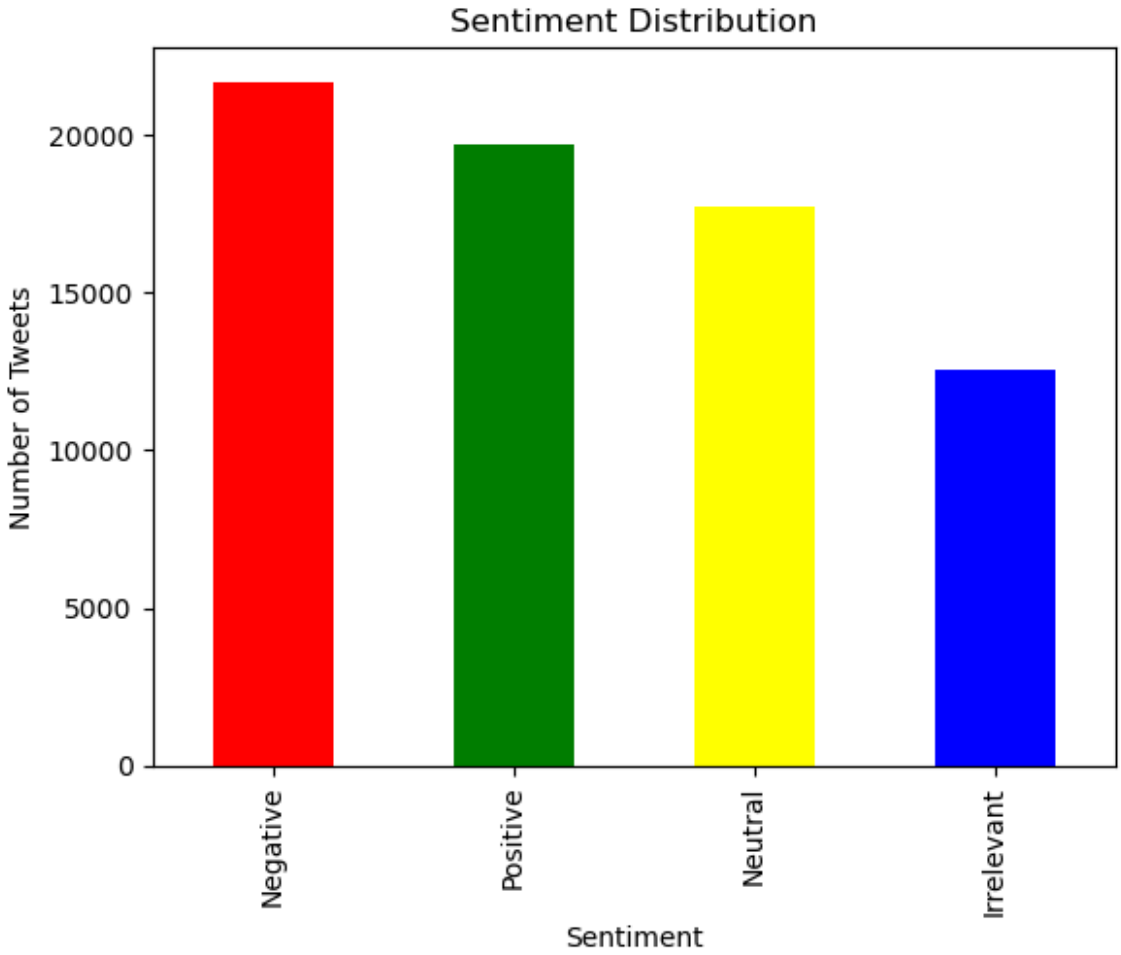
In [17]:
```
df.shape
```

Out[17]:
```
(71655, 4)
```

In [18]:
```
sentiments=df["Sentiment"].value_counts()
sentiments
```

Out[18]:
```
Sentiment
Negative      21698
Positive      19713
Neutral       17707
Irrelevant    12537
Name: count, dtype: int64
```

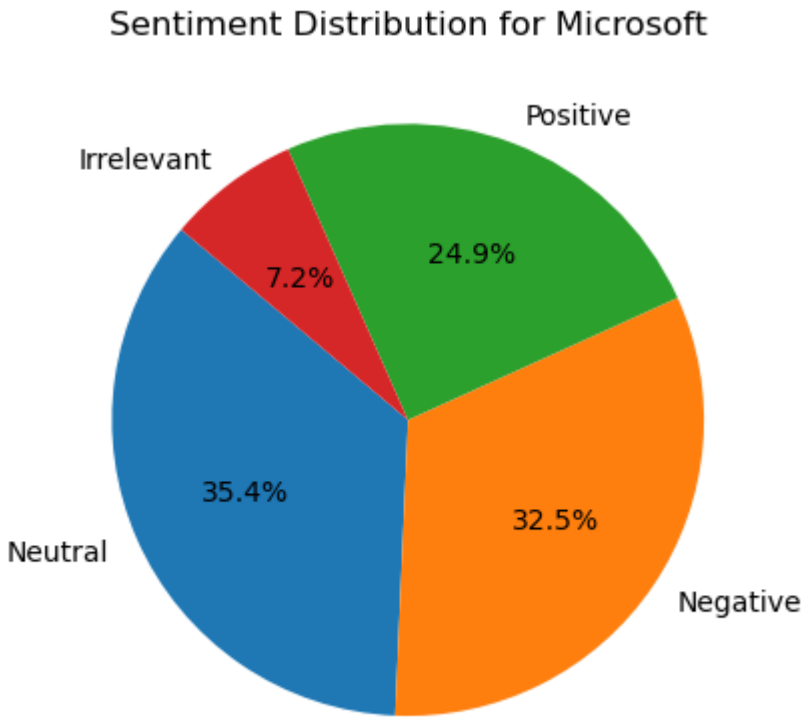
In [20]:
```
sentiments.plot(kind='bar', color=['red', 'green', 'yellow', 'blue'])
plt.title('Sentiment Distribution')
plt.xlabel('Sentiment')
plt.ylabel('Number of Tweets')
plt.show()
```



In [21]:
```
brand= df[df['Entity'].str.contains('Microsoft', case=False)]
brand_sentiment= brand['Sentiment'].value_counts()
brand_sentiment
```

Out[21]:
```
Sentiment
Neutral       816
Negative      748
Positive      573
Irrelevant    167
Name: count, dtype: int64
```

In [22]:
```
plt.pie(brand_sentiment,labels=brand_sentiment.index, autopct='%1.1f%%',sta
plt.title('Sentiment Distribution for Microsoft')
plt.show()
```



In [ ]:
```

```