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

In [14]:
```
# Load the social media data
social_media_data = pd.read_csv("C:\\Users\\rakhi\\Downloads\\archive (10)\\twitter_training.csv")
social_media_data
```

Out[14]:

	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 [15]:
```
# Display the head of the data
social_media_data.head()
```

Out[15]:

	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 [25]:
```
social_media_data.tail()
```

Out[25]:

	id	topic	sentiment	text
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...

In [26]:
```
social_media_data.describe()
```

Out[26]:

	id
count	74681.000000
mean	6432.640149
std	3740.423819
min	1.000000
25%	3195.000000
50%	6422.000000
75%	9601.000000
max	13200.000000

In [27]:
```
social_media_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 74681 entries, 0 to 74680
Data columns (total 4 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   id         74681 non-null  int64
 1   topic      74681 non-null  object
 2   sentiment  74681 non-null  object
 3   text       73995 non-null  object
dtypes: int64(1), object(3)
memory usage: 2.3+ MB
```

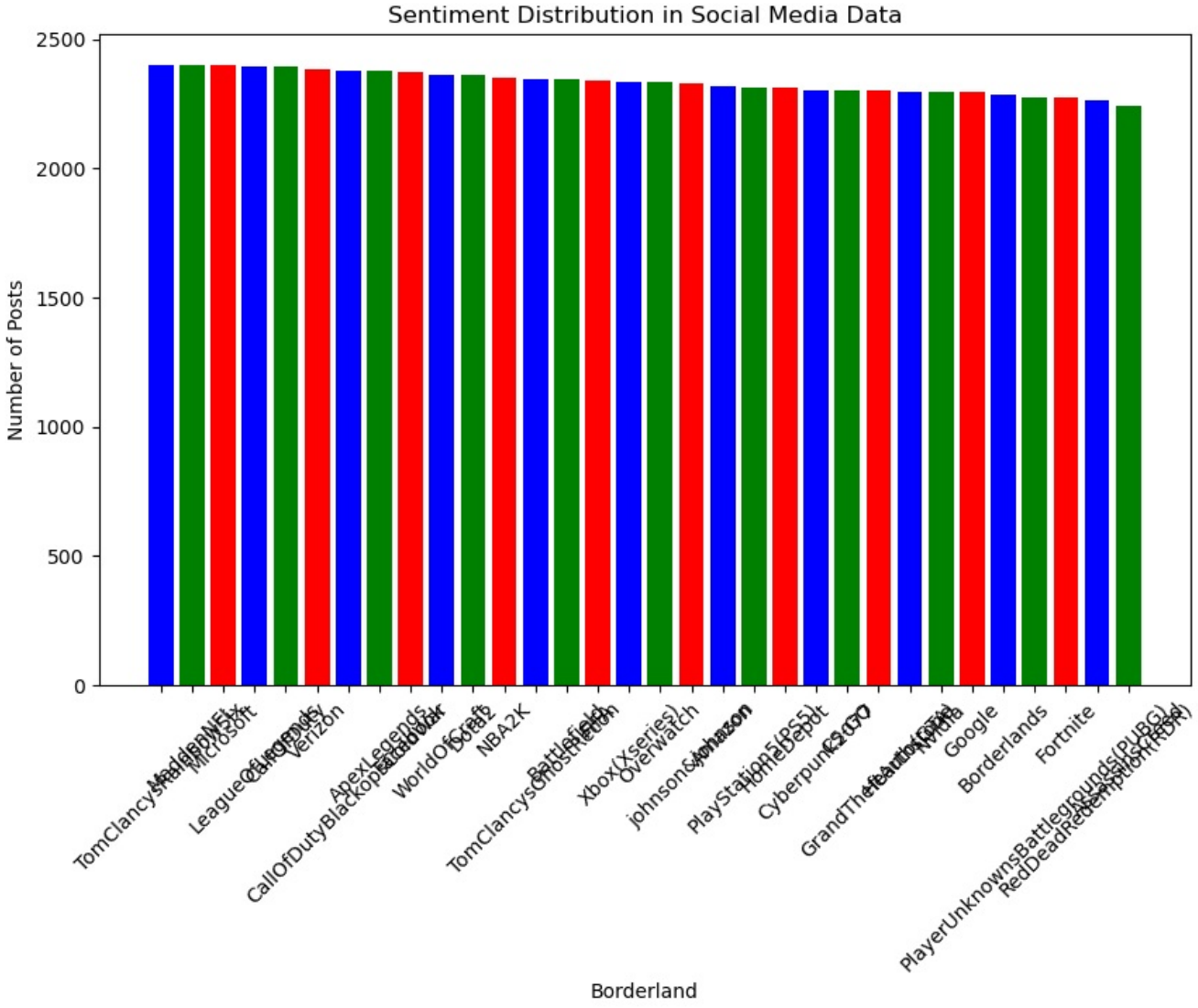
In [28]:
```
social_media_data.isnull().sum()
```

Out[28]:
```
id             0
topic          0
sentiment      0
text         686
dtype: int64
```

In [20]:
```
# Analyze the sentiment distribution
sentiment_distribution = social_media_data['Borderlands'].value_counts()
```

In [21]:
```
# Plot the sentiment distribution
plt.figure(figsize=(10, 6), facecolor='white')
plt.bar(sentiment_distribution.index, sentiment_distribution.values, color=['blue', 'green', 'red'])
plt.title('Sentiment Distribution in Social Media Data')
plt.xlabel('Borderland')
plt.ylabel('Number of Posts')
plt.xticks(rotation=45)
plt.show()

print(sentiment_distribution)
```



```
TomClancysRainbowSix                2400
MaddenNFL                           2400
Microsoft                           2400
LeagueOfLegends                     2394
CallOfDuty                          2394
Verizon                             2382
CallOfDutyBlackopsColdWar           2376
ApexLegends                         2376
Facebook                            2370

WorldOfCraft                        2364
Dota2                               2364
NBA2K                               2352
TomClancysGhostRecon                2346
Battlefield                         2346
FIFA                                2340
Xbox(Xseries)                       2334
Overwatch                           2334
johnson&johnson                     2328
Amazon                              2316
PlayStation5(PS5)                   2310
HomeDepot                           2310
Cyberpunk2077                       2304
CS-GO                               2304
GrandTheftAuto(GTA)                 2304
Hearthstone                         2298
Nvidia                              2298
Google                              2298
Borderlands                         2285
PlayerUnknownsBattlegrounds(PUBG)   2274
Fortnite                            2274
RedDeadRedemption(RDR)              2262
AssassinsCreed                      2244
Name: Borderlands, dtype: int64
```

In [18]:
```python
# Check the column names of the dataframe
column_names = social_media_data.columns
print(column_names)
```

```
Index(['2401', 'Borderlands', 'Positive',
       'im getting on borderlands and i will murder you all ,'],
      dtype='object')
```

In [23]:
```python
# Assign proper column names
columns = ['id', 'topic', 'sentiment', 'text']
social_media_data.columns = columns

# Display the corrected dataframe head to confirm
social_media_data.head()
```

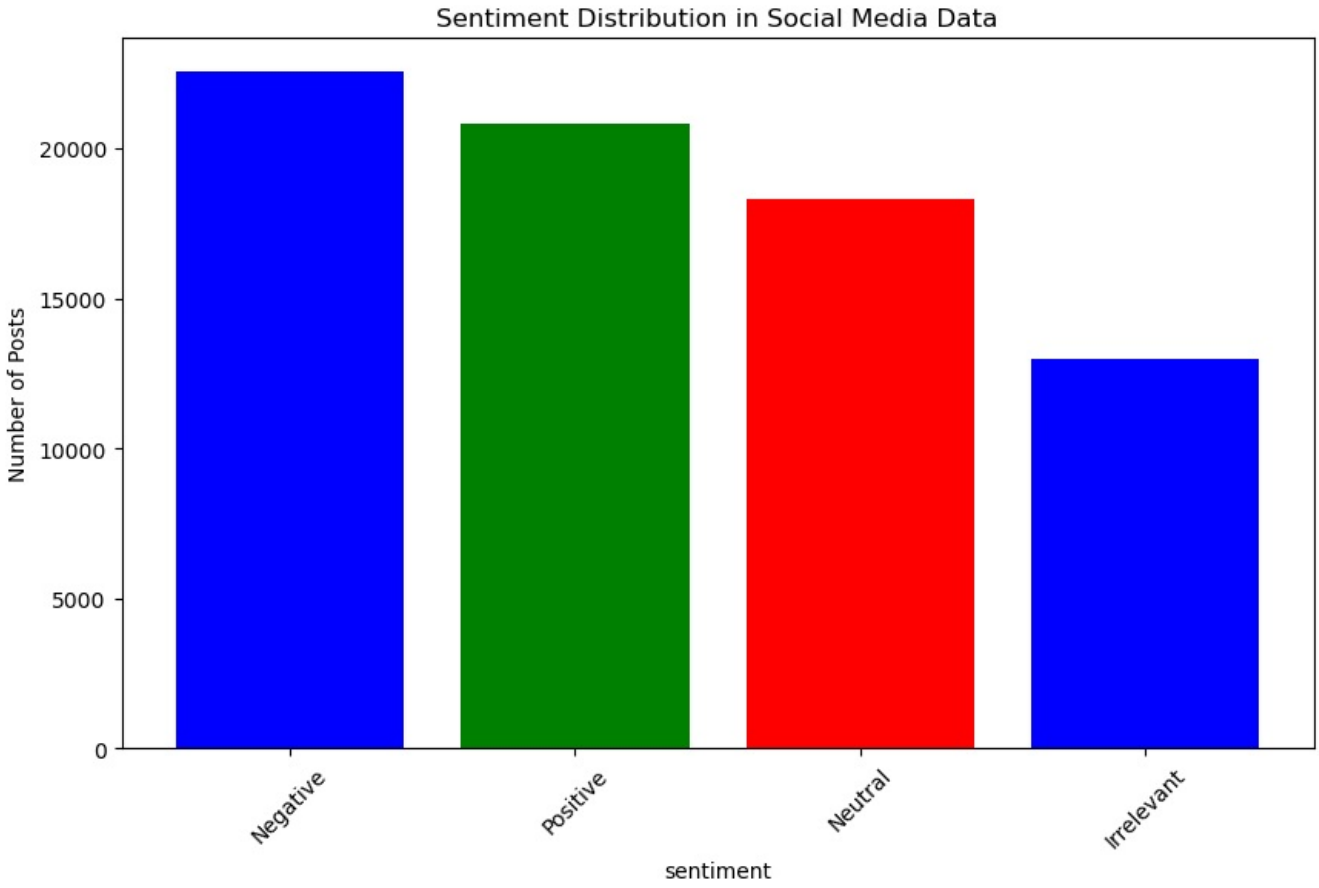
Out[23]:

	id	topic	sentiment	text
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 [24]:
```python
# Analyze the sentiment distribution
sentiment_distribution = social_media_data['sentiment'].value_counts()

# Plot the sentiment distribution
plt.figure(figsize=(10, 6), facecolor='white')
plt.bar(sentiment_distribution.index, sentiment_distribution.values, color=['blue', 'green', 'red'])
plt.title('Sentiment Distribution in Social Media Data')
plt.xlabel('sentiment')
plt.ylabel('Number of Posts')
plt.xticks(rotation=45)
plt.show()

print(sentiment_distribution)
```




```
Negative      22542
Positive      20831
Neutral       18318
Irrelevant    12990
Name: sentiment, dtype: int64
```

In [ ]: