Problem statement: Analyze and visualize sentiment patterns in social media data to understand public opinion and attitudes towards specific topics or brands Dataset: Twitter sentiment analysis About the dataset: This is an entity-level sentiment analysis dataset of twitter. Given a message and an entity, the task is to judge the sentiment of the message about the entity. There are three classes in this dataset: Positive, Negative and Neutral. We regard messages that are not relevant to the entity (i.e. Irrelevant) as Neutral.

In [ ]:
```
import pandas as pd
import matplotlib.pyplot as plt
# Load the Students performance dataset
file_path = 'c:\\Users\\Admin\\Downloads\\twitter.csv'
df = pd.read_csv(file_path)
print(df)
```

```
          textID                                               text  \
0      cb774db0d1                I`d have responded, if I were going
1      549e992a42      Sooo SAD I will miss you here in San Diego!!!
2      088c60f138                          my boss is bullying me...
3      9642c003ef                     what interview! leave me alone
4      358bd9e861   Sons of ****, why couldn`t they put them on t...
...           ...                                                ...
27476  4eac33d1c0   wish we could come see u on Denver  husband l...
27477  4f4c4fc327   I`ve wondered about rake to.  The client has ...
27478  f67aae2310   Yay good for both of you. Enjoy the break - y...
27479  ed167662a5                         But it was worth it  ****.
27480  6f7127d9d7     All this flirting going on - The ATG smiles...

                                      selected_text sentiment  \
0                    I`d have responded, if I were going   neutral
1                                          Sooo SAD  negative
2                                        bullying me  negative
3                                      leave me alone  negative
4                                       Sons of ****,  negative
...                                               ...       ...
27476                                        d lost  negative
27477                                     , don`t force  negative
27478                          Yay good for both of you.  positive
27479                         But it was worth it  ****.  positive
27480  All this flirting going on - The ATG smiles. Y...   neutral

              entity
0         Borderlands
1         Borderlands
2         Borderlands
3         Borderlands
4         Borderlands
...               ...
27476  AssassinsCreed
27477  AssassinsCreed
27478  AssassinsCreed
27479  AssassinsCreed
27480  AssassinsCreed

[27481 rows x 5 columns]
```

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

Out[ ]:
```
(27481, 5)
```

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

Out[ ]:
```
textID          0
text            1
selected_text    1
sentiment        0
entity           0
dtype: int64
```

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

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

Out[ ]:
```
textID          0
text            0
selected_text    0
sentiment        0
entity           0
dtype: int64
```

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

Out[ ]:
```
0
```

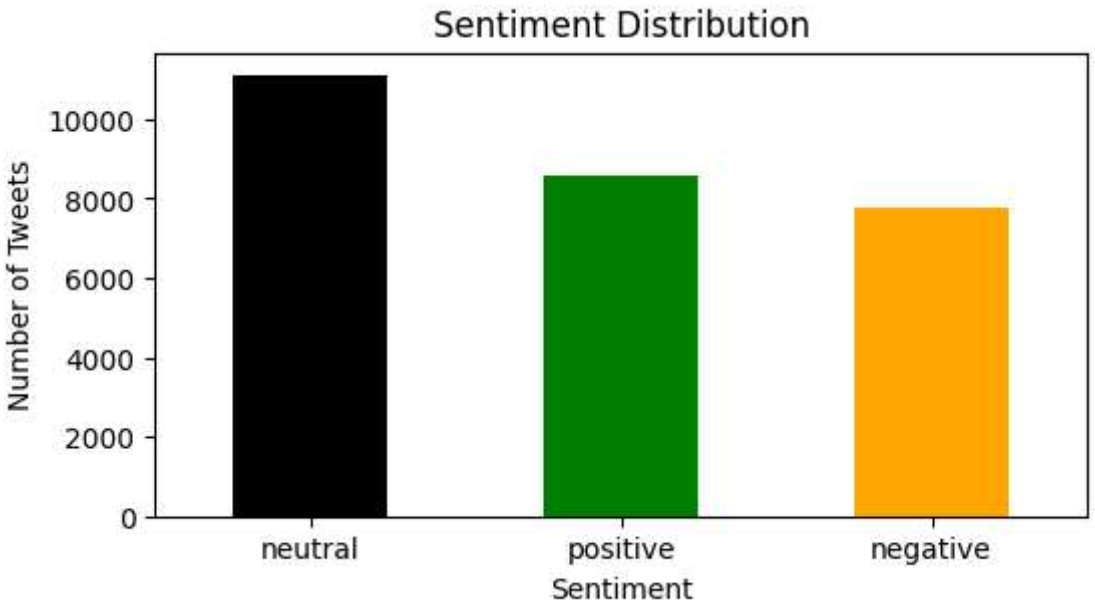
In [ ]:
```
sentiment_counts = df['sentiment'].value_counts()
sentiment_counts
```

Out[ ]:
```
sentiment
neutral     11117
positive     8582
negative     7781
Name: count, dtype: int64
```

In [ ]:
```
df.columns
```

Out[ ]:
```
Index(['textID', 'text', 'selected_text', 'sentiment', 'entity'], dtype='object')
```

In [ ]:
```
plt.figure(figsize=(6, 3))
sentiment_counts.plot(kind='bar', color=['black', 'green', 'orange', 'brown'])
plt.title('Sentiment Distribution')
plt.xlabel('Sentiment')
plt.ylabel('Number of Tweets')
plt.xticks(rotation=0)
plt.show()
```



In [ ]:
```
df['entity'].value_counts()
```

Out[ ]:
```
entity
CallOfDutyBlackopsColdWar    2376
Dota2                        2364
WorldOfCraft                 2364
NBA2K                        2352
Overwatch                    2334
Xbox(Xseries)                2334
Amazon                       2316
PlayStation5(PS5)            2310
CS-GO                        2304
Google                       2298
Borderlands                  2285
AssassinsCreed               1189
ApexLegends                   654
Name: count, dtype: int64
```

In [ ]:
```
brand_data = df[df['entity'].str.contains('Google', case=False)]
brand_sentiment_counts = brand_data['sentiment'].value_counts()
brand_sentiment_counts
```

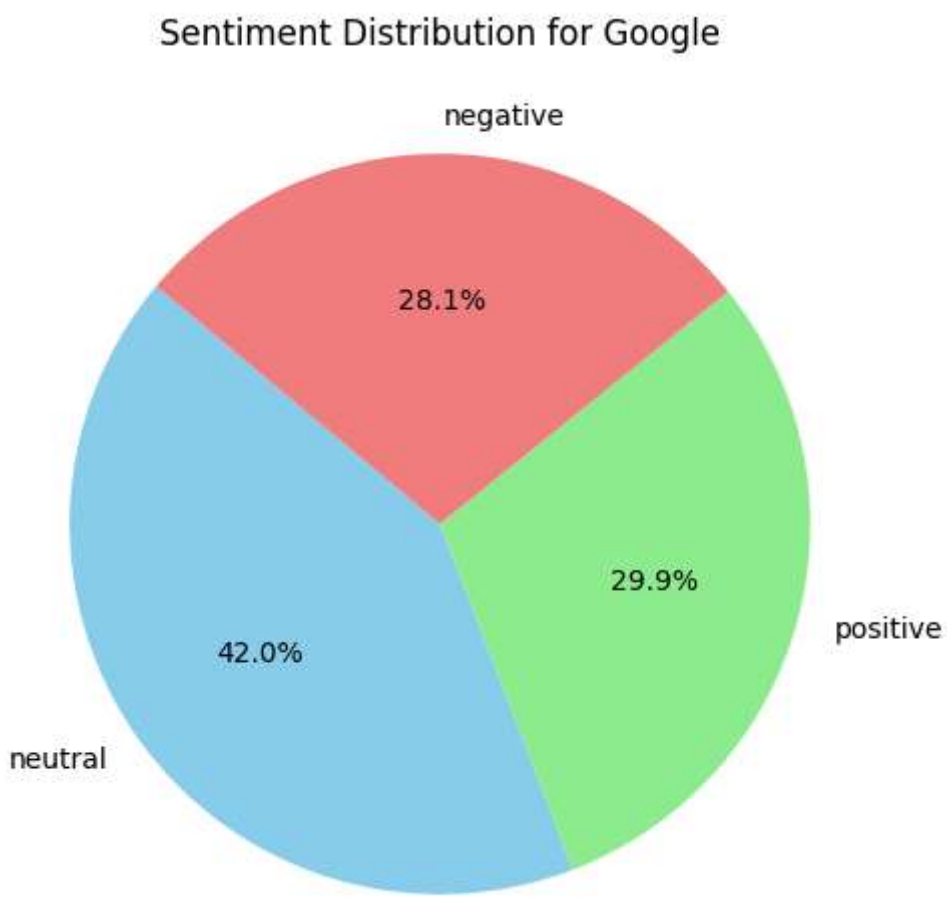
Out[ ]:
```
sentiment
neutral     945
positive    696
negative    657
Name: count, dtype: int64
```

In [ ]:
```
# Define a custom color palette
custom_colors = ['skyblue', 'lightgreen', 'lightcoral', 'Pink']

plt.figure(figsize=(6, 6))

# Assuming brand_sentiment_counts is a pandas Series containing sentiment counts
plt.pie(brand_sentiment_counts, labels=brand_sentiment_counts.index, autopct='%1.1f

plt.title('Sentiment Distribution for Google')
plt.show()
```



In [ ]:
```
brand_data = df[df['entity'].str.contains('Dota2', case=False)]
brand_sentiment_counts = brand_data['sentiment'].value_counts()
brand_sentiment_counts
```

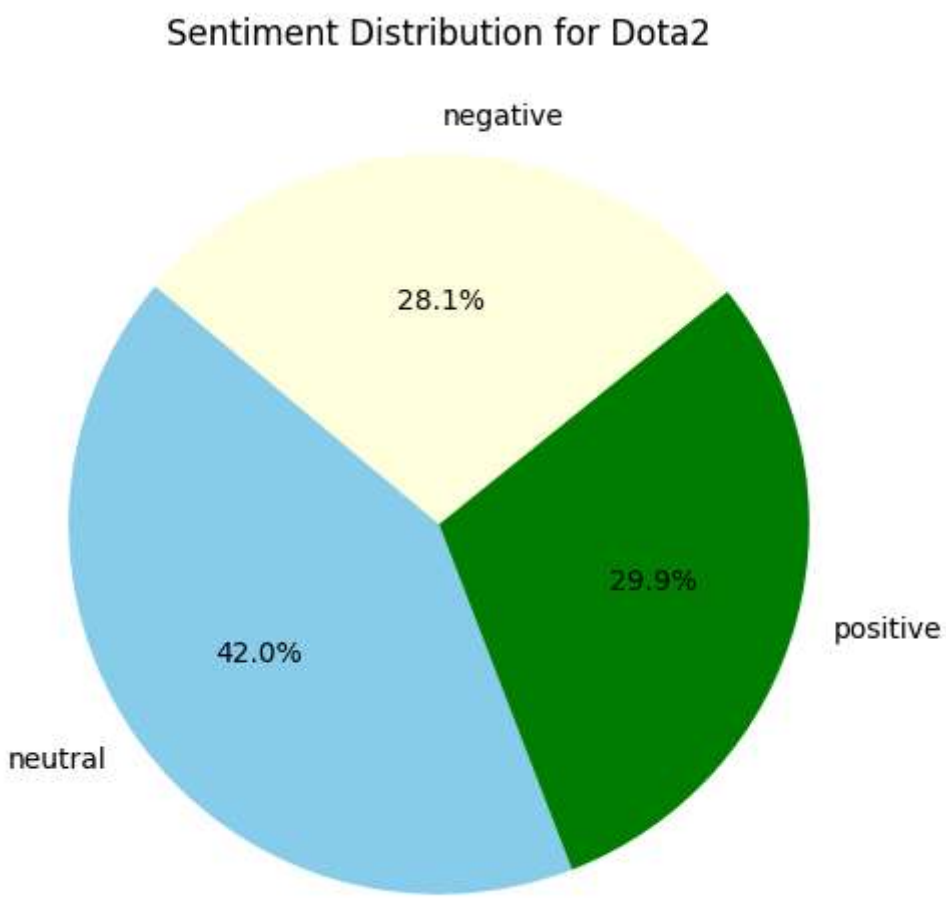
Out[ ]:
```
sentiment
neutral     992
positive    708
negative    664
Name: count, dtype: int64
```

In [ ]:
```
# Define a custom color palette
custom_colors = ['skyblue', 'green', 'lightyellow', 'red']

plt.figure(figsize=(6, 6))

# Assuming brand_sentiment_counts is a pandas Series containing sentiment counts
plt.pie(brand_sentiment_counts, labels=brand_sentiment_counts.index, autopct='%1.1f

plt.title('Sentiment Distribution for Dota2')
plt.show()
```


Sentiment Distribution for Dota2
negative
28.1%
29.9%
positive
42.0%
neutral

In [ ]:

In [ ]: