In [4]: `!pip install wordcloud`

```
Collecting wordcloud
  Downloading wordcloud-1.9.4-cp312-cp312-macosx_11_0_arm64.whl.metadata (3.4 kB)
Requirement already satisfied: numpy>=1.6.1 in /opt/anaconda3/lib/python3.12/site-packages (from wordcloud) (1.26.4)
Requirement already satisfied: pillow in /opt/anaconda3/lib/python3.12/site-packages (from wordcloud) (10.4.0)
Requirement already satisfied: matplotlib in /opt/anaconda3/lib/python3.12/site-packages (from wordcloud) (3.9.2)
Requirement already satisfied: contourpy>=1.0.1 in /opt/anaconda3/lib/python3.12/site-packages (from matplotlib->wor
dcloud) (1.2.0)
Requirement already satisfied: cycler>=0.10 in /opt/anaconda3/lib/python3.12/site-packages (from matplotlib->wordclo
ud) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in /opt/anaconda3/lib/python3.12/site-packages (from matplotlib->wo
rdcloud) (4.51.0)
Requirement already satisfied: kiwisolver>=1.3.1 in /opt/anaconda3/lib/python3.12/site-packages (from matplotlib->wo
rdcloud) (1.4.4)
Requirement already satisfied: packaging>=20.0 in /opt/anaconda3/lib/python3.12/site-packages (from matplotlib->word
cloud) (24.1)
Requirement already satisfied: pyparsing>=2.3.1 in /opt/anaconda3/lib/python3.12/site-packages (from matplotlib->wor
dcloud) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in /opt/anaconda3/lib/python3.12/site-packages (from matplotlib-
>wordcloud) (2.9.0.post0)
Requirement already satisfied: six>=1.5 in /opt/anaconda3/lib/python3.12/site-packages (from python-dateutil>=2.7->m
atplotlib->wordcloud) (1.16.0)
Downloading wordcloud-1.9.4-cp312-cp312-macosx_11_0_arm64.whl (168 kB)
Installing collected packages: wordcloud
Successfully installed wordcloud-1.9.4
```

In [8]:
```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from wordcloud import WordCloud

df = pd.read_excel("twitter_training.xlsx")

print(df.info())
print(df['sentiment'].value_counts())

df = df.drop_duplicates()
```

```python
plt.figure(figsize=(6, 4))
sns.countplot(data=df, x='sentiment', hue='sentiment', palette='Set2', legend=False)
plt.title("Overall Sentiment Distribution")
plt.xlabel("Sentiment")
plt.ylabel("Tweet Count")
plt.tight_layout()
plt.show()

plt.figure(figsize=(10, 6))
sns.countplot(data=df, x='entity', hue='sentiment', palette='Set1')
plt.title("Sentiment Count by Entity (Brand/Topic)")
plt.xticks(rotation=90)
plt.tight_layout()
plt.show()

for sentiment in df['sentiment'].unique():
    text = " ".join(df[df['sentiment'] == sentiment]['tweet'].astype(str))
    wc = WordCloud(width=800, height=400, background_color='white').generate(text)
    plt.figure(figsize=(10, 5))
    plt.imshow(wc, interpolation='bilinear')
    plt.axis('off')
    plt.title(f"Word Cloud for {sentiment} Tweets")
    plt.show()

heatmap_data = pd.crosstab(df['entity'], df['sentiment'])
plt.figure(figsize=(12, 6))
sns.heatmap(heatmap_data, annot=True, fmt='d', cmap='YlGnBu')
plt.title("Heatmap of Sentiments per Entity")
plt.ylabel("Entity")
plt.xlabel("Sentiment")
plt.tight_layout()
plt.show()

print("\nTop 5 Positive Tweets:")
print(df[df['sentiment'] == 'Positive']['tweet'].head(5), "\n")

print("Top 5 Negative Tweets:")
print(df[df['sentiment'] == 'Negative']['tweet'].head(5))
```