

```

# task 4

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from wordcloud import WordCloud
from textblob import TextBlob

# Read file with space as delimiter and no header
df = pd.read_csv("TWIT.csv", sep=None, engine='python', header=None)

# Combine the last columns into a full "text" string (from 3rd column onward)
df["id"] = df[0]
df["topic"] = df[1]
df["original_sentiment"] = df[2]
df["text"] = df.iloc[:, 3:].astype(str).agg(" ".join, axis=1)

# Keep only relevant columns
df = df[["id", "topic", "original_sentiment", "text"]]

# Show sample
print("Sample data:")
print(df.head())

# Add sentiment using TextBlob
def get_sentiment(text):
    blob = TextBlob(str(text))
    polarity = blob.sentiment.polarity
    if polarity > 0:
        return 'Positive'
    elif polarity < 0:
        return 'Negative'

```

```
else:
```

```
    return 'Neutral'
```

```
df["Sentiment"] = df["text"].apply(get_sentiment)
```

```
# Sentiment distribution
```

```
print("\nSentiment Counts:")
```

```
print(df["Sentiment"].value_counts())
```

```
# Plot sentiment distribution
```

```
plt.figure(figsize=(6, 4))
```

```
sns.countplot(data=df, x="Sentiment", palette="Set2")
```

```
plt.title("Sentiment Distribution")
```

```
plt.savefig("sentiment_distribution.png")
```

```
plt.show()
```

```
# Word cloud for positive tweets
```

```
positive_tweets = " ".join(df[df["Sentiment"] == "Positive"]["text"].dropna())
```

```
wordcloud = WordCloud(width=800, height=400,  
background_color='white').generate(positive_tweets)
```

```
plt.figure(figsize=(10, 5))
```

```
plt.imshow(wordcloud, interpolation="bilinear")
```

```
plt.axis("off")
```

```
plt.title("WordCloud - Positive Tweets")
```

```
plt.savefig("positive_wordcloud.png")
```

```
plt.show()
```