In [9]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import re
import string
import nltk
import warnings
%matplotlib inline

warnings.filterwarnings('ignore')
```

In [14]:

```
df=pd.read_csv('twitter_validation.csv')
df.head()
```

Out[14]:

	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by Tom’s great auntie as ‘Hayley can’t get out of bed’ and told to his grandma, who now thinks I’m a lazy, terrible person 🤣
0	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
2	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
3	4433	Google	Neutral	Now the President is slapping Americans in the...
4	6273	FIFA	Negative	Hi @EAHelp I’ve had Madeleine McCann in my cel...

In [15]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 999 entries, 0 to 998
Data columns (total 4 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   3364
999 non-null    int64
 1   Facebook
999 non-null    object
 2   Irrelevant
999 non-null    object
 3   I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by To
m's great auntie as 'Hayley can't get out of bed' and told to his grandma, who now thinks I'm a lazy, terrible person 🤣  999 no
n-null    object
dtypes: int64(1), object(3)
memory usage: 31.3+ KB
```

In [ ]:

```

```

In [ ]:

```python
def remove_pattern(input_txt, pattern):
    r=re.findall(pattern,input_txt)
    for word in r:
        input_txt=re.sub(word, "",input_txt)
```

In [ ]:

```python
df['clean_tweet']=np.vectorize(remove_pattern)(df['tweet'],"@[\w]*")
```

In [20]:

```python
df.head()
```

Out[20]:

	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by Tom's great auntie as 'Hayley can't get out of bed' and told to his grandma, who now thinks I'm a lazy, terrible person 🤣
0	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
2	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
3	4433	Google	Neutral	Now the President is slapping Americans in the...
4	6273	FIFA	Negative	Hi @EAHelp I've had Madeleine McCann in my cel...

In [18]:

```
def remove_pattern(input_txt, pattern):
    r=re.findall(pattern,input_txt)
    for word in r:
        input_txt=re.sub(word, "",input_txt)
        df['clean_tweet']=np.vectorize(remove_pattern)(df['tweet'],"@[\w]*")
df.head()
```

Out[18]:

	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by Tom’s great auntie as ‘Hayley can’t get out of bed’ and told to his grandma, who now thinks I’m a lazy, terrible person 🤣
0	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
2	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
3	4433	Google	Neutral	Now the President is slapping Americans in the...
4	6273	FIFA	Negative	Hi @EAHelp I’ve had Madeleine McCann in my cel...

In [ ]:

```
def remove_pattern(input_txt, pattern):
    r=re.findall(pattern,input_txt)
    for word in r:
        input_txt=re.sub(word, "",input_txt)
```

In [21]:

```
df.head()
```

Out[21]:

	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by Tom’s great auntie as ‘Hayley can’t get out of bed’ and told to his grandma, who now thinks I’m a lazy, terrible person 🤣
0	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
2	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
3	4433	Google	Neutral	Now the President is slapping Americans in the...
4	6273	FIFA	Negative	Hi @EAHelp I’ve had Madeleine McCann in my cel...

In [ ]:

```
df['clean_tweet']=np.vectorize(remove_pattern)(df['tweet'], "@[\w]*")
```

In [22]:
```
df.head()
```

Out[22]:

	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by Tom’s great auntie as ‘Hayley can’t get out of bed’ and told to his grandma, who now thinks I’m a lazy, terrible person 🤣
0	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
2	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
3	4433	Google	Neutral	Now the President is slapping Americans in the...
4	6273	FIFA	Negative	Hi @EAHelp I’ve had Madeleine McCann in my cel...

In [27]:
```
def remove_pattern(df):
    cleaned_data=df.replace('@','')
    return cleaned_data
print(cleaned_data)
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Cell In[27], line 4
      2     cleaned_data=df.replace('@','')
      3     return cleaned_data
----> 4 print(cleaned_data)

NameError: name 'cleaned_data' is not defined
```

In [28]:
```
import pandas as pd

df=pd.read_csv('twitter_validation.csv')
df.head()
```

Out[28]:

	3364	Facebook	Irrelevant	I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by Tom’s great auntie as ‘Hayley can’t get out of bed’ and told to his grandma, who now thinks I’m a lazy, terrible person 🤣
0	352	Amazon	Neutral	BBC News - Amazon boss Jeff Bezos rejects clai...
1	8312	Microsoft	Negative	@Microsoft Why do I pay for WORD when it funct...
2	4371	CS-GO	Negative	CSGO matchmaking is so full of closet hacking,...
3	4433	Google	Neutral	Now the President is slapping Americans in the...
4	6273	FIFA	Negative	Hi @EAHelp I’ve had Madeleine McCann in my cel...

In [29]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 999 entries, 0 to 998
Data columns (total 4 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   3364
999 non-null    int64
 1   Facebook
999 non-null    object
 2   Irrelevant
999 non-null    object
 3   I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated by To
m’s great auntie as ‘Hayley can’t get out of bed’ and told to his grandma, who now thinks I’m a lazy, terrible person 🤣  999 no
n-null    object
dtypes: int64(1), object(3)
memory usage: 31.3+ KB
```

In [30]:

```
df.describe()
```

Out[30]:

	3364
count	999.000000
mean	6435.159159
std	3728.912226
min	6.000000
25%	3241.500000
50%	6560.000000
75%	9662.500000
max	13197.000000

In [35]:

```python
#sentiment distribution
import seaborn as sns
import matplotlib.pyplot as plt

sns.countplot(x='Irrelevant', df=df)
plt.title('sentiment distribution')
print(sns.countplot)
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[35], line 5
      2 import seaborn as sns
      3 import matplotlib.pyplot as plt
----> 5 sns.countplot(x='Irrelevant', df=df)
      6 plt.title('sentiment distribution')
      7 print(sns.countplot)

File ~\anaconda3\Lib\site-packages\seaborn\categorical.py:2943, in countplot(data, x, y, hue, order, hue_order, orient, color, p
alette, saturation, width, dodge, ax, **kwargs)
   2940 elif x is not None and y is not None:
   2941     raise ValueError("Cannot pass values for both `x` and `y`")
-> 2943 plotter = _CountPlotter(
   2944     x, y, hue, data, order, hue_order,
   2945     estimator, errorbar, n_boot, units, seed,
   2946     orient, color, palette, saturation,
   2947     width, errcolor, errwidth, capsize, dodge
   2948 )
   2950 plotter.value_label = "count"
   2952 if ax is None:

File ~\anaconda3\Lib\site-packages\seaborn\categorical.py:1530, in _BarPlotter.__init__(self, x, y, hue, data, order, hue_order,
estimator, errorbar, n_boot, units, seed, orient, color, palette, saturation, width, errcolor, errwidth, capsize, dodge)
   1525 def __init__(self, x, y, hue, data, order, hue_order,
   1526              estimator, errorbar, n_boot, units, seed,
   1527              orient, color, palette, saturation, width,
   1528              errcolor, errwidth, capsize, dodge):
   1529     """Initialize the plotter."""
-> 1530     self.establish_variables(x, y, hue, data, orient,
   1531                              order, hue_order, units)
   1532     self.establish_colors(color, palette, saturation)
   1533     self.estimate_statistic(estimator, errorbar, n_boot, seed)

File ~\anaconda3\Lib\site-packages\seaborn\categorical.py:541, in _CategoricalPlotter.establish_variables(self, x, y, hue, data,
orient, order, hue_order, units)
    539     if isinstance(var, str):
    540         err = f"Could not interpret input '{var}'"
--> 541         raise ValueError(err)
    543 # Figure out the plotting orientation
    544 orient = infer_orient(
    545     x, y, orient, require_numeric=self.require_numeric
    546 )
```

```
ValueError: Could not interpret input 'Irrelevant'
```

In [36]:

```
df['tweet_length'] = df['text'].apply(len)
sns.histplot(df['tweet_length'])
plt.title('Distribution of Tweet Lengths')
plt.show()
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3653, in Index.get_loc(self, key)
   3652 try:
-> 3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:147, in pandas._libs.index.IndexEngine.get_loc()

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:176, in pandas._libs.index.IndexEngine.get_loc()

File pandas\_libs\hashtable_class_helper.pxi:7080, in pandas._libs.hashtable.PyObjectHashTable.get_item()

File pandas\_libs\hashtable_class_helper.pxi:7088, in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: 'text'

The above exception was the direct cause of the following exception:

KeyError                                  Traceback (most recent call last)
Cell In[36], line 1
----> 1 df['tweet_length'] = df['text'].apply(len)
      2 sns.histplot(df['tweet_length'])
      3 plt.title('Distribution of Tweet Lengths')

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:3761, in DataFrame.__getitem__(self, key)
   3759 if self.columns.nlevels > 1:
   3760     return self._getitem_multilevel(key)
-> 3761 indexer = self.columns.get_loc(key)
   3762 if is_integer(indexer):
   3763     indexer = [indexer]

File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3655, in Index.get_loc(self, key)
   3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:
-> 3655     raise KeyError(key) from err
   3656 except TypeError:
   3657     # If we have a listlike key, _check_indexing_error will raise
   3658     #  InvalidIndexError. Otherwise we fall through and re-raise
   3659     #  the TypeError.
   3660     self._check_indexing_error(key)

KeyError: 'text'
```

In [39]:

```
from wordcloud import WordCloud

# Create a function to generate word clouds
def generate_wordcloud(text, title):
    wordcloud = WordCloud(width=800, height=400, random_state=42, max_words=100).generate(text)
    plt.figure(figsize=(10, 5))
    plt.imshow(wordcloud, interpolation='bilinear')
    plt.axis('off')
    plt.title(title)
    plt.show()

# Example usage:
positive_tweets = data[data['sentiment'] == 'positive']['text'].str.cat(sep=' ')
generate_wordcloud(positive_tweets, 'Positive Tweets Word Cloud')
```

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
Cell In[39], line 1
----> 1 from wordcloud import WordCloud
      3 # Create a function to generate word clouds
      4 def generate_wordcloud(text, title):

ModuleNotFoundError: No module named 'wordcloud'
```

In [41]:

```
def extract_hashtags(text):
    return [word.lower() for word in text.split() if word.startswith('#')]

df['hashtags'] = df['str'].apply(extract_hashtags)
hashtags = [hashtag for sublist in df['hashtags'] for hashtag in sublist]

# Plot the top N hashtags
top_hashtags = pd.Series(hashtags).value_counts().nlargest(10)
top_hashtags.plot(kind='bar', color='skyblue')
plt.title('Top 10 Hashtags in Tweets')
plt.xlabel('Hashtags')
plt.ylabel('Frequency')
plt.show()
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3653, in Index.get_loc(self, key)
   3652 try:
-> 3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:147, in pandas._libs.index.IndexEngine.get_loc()

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:176, in pandas._libs.index.IndexEngine.get_loc()

File pandas\_libs\hashtable_class_helper.pxi:7080, in pandas._libs.hashtable.PyObjectHashTable.get_item()

File pandas\_libs\hashtable_class_helper.pxi:7088, in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: 'str'

The above exception was the direct cause of the following exception:

KeyError                                  Traceback (most recent call last)
Cell In[41], line 4
      1 def extract_hashtags(text):
      2     return [word.lower() for word in text.split() if word.startswith('#')]
----> 4 df['hashtags'] = df['str'].apply(extract_hashtags)
      5 hashtags = [hashtag for sublist in df['hashtags'] for hashtag in sublist]
      7 # Plot the top N hashtags

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:3761, in DataFrame.__getitem__(self, key)
   3759 if self.columns.nlevels > 1:
   3760     return self._getitem_multilevel(key)
-> 3761 indexer = self.columns.get_loc(key)
   3762 if is_integer(indexer):
   3763     indexer = [indexer]

File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3655, in Index.get_loc(self, key)
   3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:
-> 3655     raise KeyError(key) from err
   3656 except TypeError:
   3657     # If we have a listlike key, _check_indexing_error will raise
   3658     #  InvalidIndexError. Otherwise we fall through and re-raise
   3659     #  the TypeError.
   3660     self._check_indexing_error(key)
```

```
KeyError: 'str'
```

In [45]:

```
import nltk
from nltk.sentiment import SentimentIntensityAnalyzer

nltk.download('vader_lexicon')

def analyze_sentiment(text):
    sid = SentimentIntensityAnalyzer()
    sentiment_score = sid.polarity_scores(text)['compound']

    if sentiment_score >= 0.05:
        return 'Positive'
    elif sentiment_score <= -0.05:
        return 'Negative'
    else:
        return 'Neutral'

if __name__ == "__main__":
    # Example text for sentiment analysis
    text = "I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translated

    sentiment = analyze_sentiment(text)
    print(f"Sentiment: {sentiment}")
```

```
Sentiment: Negative
```

```
[nltk_data] Downloading package vader_lexicon to
[nltk_data]     C:\Users\HP\AppData\Roaming\nltk_data...
[nltk_data]   Package vader_lexicon is already up-to-date!
```

In [46]:

```
import pandas as pd
import matplotlib.pyplot as plt


df = pd.read_csv('twitter_validation.csv')


df['tweet_length'] = df['text'].apply(len)


average_length_by_sentiment = df.groupby('sentiment')['tweet_length'].mean()
```

```
plt.figure(figsize=(10, 6))
average_length_by_sentiment.plot(kind='bar', color=['green', 'yellow', 'red'])
plt.title('Average Tweet Length by Sentiment')
plt.xlabel('Sentiment')
plt.ylabel('Average Tweet Length')
plt.xticks(rotation=0)
plt.show()
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3653, in Index.get_loc(self, key)
   3652 try:
-> 3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:147, in pandas._libs.index.IndexEngine.get_loc()

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:176, in pandas._libs.index.IndexEngine.get_loc()

File pandas\_libs\hashtable_class_helper.pxi:7080, in pandas._libs.hashtable.PyObjectHashTable.get_item()

File pandas\_libs\hashtable_class_helper.pxi:7088, in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: 'text'

The above exception was the direct cause of the following exception:

KeyError                                  Traceback (most recent call last)
Cell In[46], line 8
      2 import matplotlib.pyplot as plt
      5 df = pd.read_csv('twitter_validation.csv')
----> 8 df['tweet_length'] = df['text'].apply(len)
     11 average_length_by_sentiment = df.groupby('sentiment')['tweet_length'].mean()
     14 plt.figure(figsize=(10, 6))

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:3761, in DataFrame.__getitem__(self, key)
   3759 if self.columns.nlevels > 1:
   3760     return self._getitem_multilevel(key)
-> 3761 indexer = self.columns.get_loc(key)
   3762 if is_integer(indexer):
   3763     indexer = [indexer]

File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3655, in Index.get_loc(self, key)
   3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:
-> 3655     raise KeyError(key) from err
   3656 except TypeError:
   3657     # If we have a listlike key, _check_indexing_error will raise
   3658     #  InvalidIndexError. Otherwise we fall through and re-raise
   3659     #  the TypeError.
   3660     self._check_indexing_error(key)
```

```
KeyError: 'text'
```

In [47]:

```
tweet = data()
words = tweet.split()
word_count = len(words)
print(f"The length of the tweet in words: {word_count}")
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
Cell In[47], line 2
      1 data=('twitter_validation.csv')
----> 2 tweet = data()
      3 words = tweet.split()
      4 word_count = len(words)

TypeError: 'str' object is not callable
```

In [49]:

```
!pip install wordcloud
```

```
Requirement already satisfied: wordcloud in c:\users\hp\anaconda3\lib\site-packages (1.9.3)
Requirement already satisfied: numpy>=1.6.1 in c:\users\hp\anaconda3\lib\site-packages (from wordcloud) (1.24.3)
Requirement already satisfied: pillow in c:\users\hp\anaconda3\lib\site-packages (from wordcloud) (9.4.0)
Requirement already satisfied: matplotlib in c:\users\hp\anaconda3\lib\site-packages (from wordcloud) (3.7.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (4.25.
0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (23.1)
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (3.
0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (2.
8.2)
Requirement already satisfied: six>=1.5 in c:\users\hp\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib->wordc
loud) (1.16.0)
```

In [54]:

```
import pandas as pd
from wordcloud import WordCloud
import matplotlib.pyplot as plt

# Load your dataset
# Replace 'your_dataset.csv' with the actual file name
df = pd.read_csv('twitter_validation.csv')
```

```
# Concatenate all text from the 'text' column
all_text = ' '.join(df['twitter_validation.csv'].dropna())

# Generate a word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(all_text)

# Display the generated word cloud using Matplotlib
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')  # Turn off the axis
plt.show()
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3653, in Index.get_loc(self, key)
   3652 try:
-> 3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:147, in pandas._libs.index.IndexEngine.get_loc()

File ~\anaconda3\Lib\site-packages\pandas\_libs\index.pyx:176, in pandas._libs.index.IndexEngine.get_loc()

File pandas\_libs\hashtable_class_helper.pxi:7080, in pandas._libs.hashtable.PyObjectHashTable.get_item()

File pandas\_libs\hashtable_class_helper.pxi:7088, in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: 'twitter_validation.csv'

The above exception was the direct cause of the following exception:

KeyError                                  Traceback (most recent call last)
Cell In[54], line 10
      7 df = pd.read_csv('twitter_validation.csv')
      9 # Concatenate all text from the 'text' column
---> 10 all_text = ' '.join(df['twitter_validation.csv'].dropna())
     12 # Generate a word cloud
     13 wordcloud = WordCloud(width=800, height=400, background_color='white').generate(all_text)

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:3761, in DataFrame.__getitem__(self, key)
   3759 if self.columns.nlevels > 1:
   3760     return self._getitem_multilevel(key)
-> 3761 indexer = self.columns.get_loc(key)
   3762 if is_integer(indexer):
   3763     indexer = [indexer]

File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3655, in Index.get_loc(self, key)
   3653     return self._engine.get_loc(casted_key)
   3654 except KeyError as err:
-> 3655     raise KeyError(key) from err
   3656 except TypeError:
   3657     # If we have a listlike key, _check_indexing_error will raise
   3658     #  InvalidIndexError. Otherwise we fall through and re-raise
   3659     #  the TypeError.
   3660     self._check_indexing_error(key)
```

```
KeyError: 'twitter_validation.csv'
```

In [56]:

```
!pip install wordcloud
```

```
Requirement already satisfied: wordcloud in c:\users\hp\anaconda3\lib\site-packages (1.9.3)
Requirement already satisfied: numpy>=1.6.1 in c:\users\hp\anaconda3\lib\site-packages (from wordcloud) (1.24.3)
Requirement already satisfied: pillow in c:\users\hp\anaconda3\lib\site-packages (from wordcloud) (9.4.0)
Requirement already satisfied: matplotlib in c:\users\hp\anaconda3\lib\site-packages (from wordcloud) (3.7.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (4.25.
0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (23.1)
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (3.
0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\hp\anaconda3\lib\site-packages (from matplotlib->wordcloud) (2.
8.2)
Requirement already satisfied: six>=1.5 in c:\users\hp\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib->wordc
loud) (1.16.0)
```

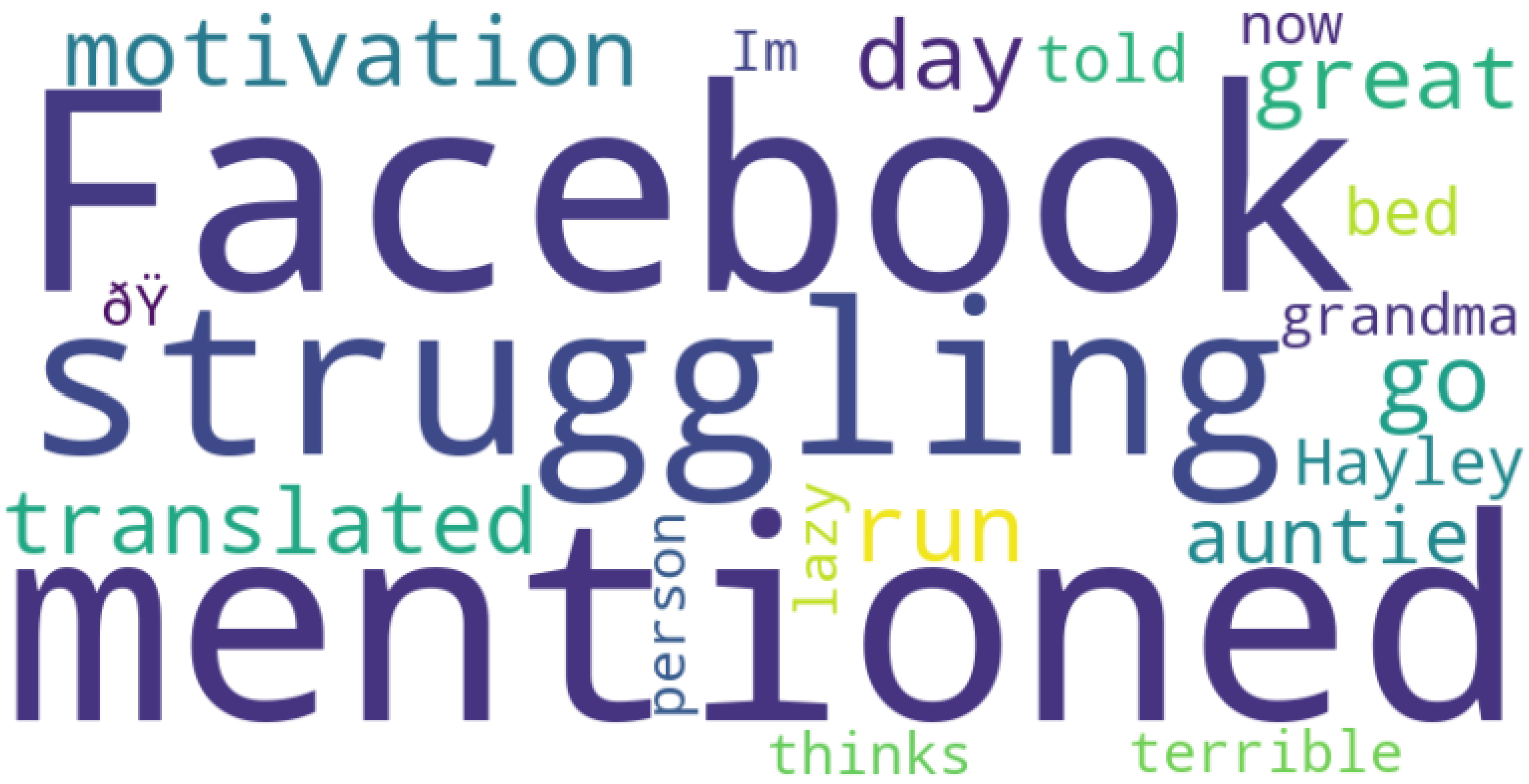
In [61]:

```
import pandas as pd
from wordcloud import WordCloud
import matplotlib.pyplot as plt

# Replace this with your own dataset or text data
df = pd.read_csv('twitter_validation.csv')
#all_text = ' '.join(df['Facebook'].astype(str))
text_data = "I mentioned on Facebook that I was struggling for motivation to go for a run the other day, which has been translate

# Generate a word cloud
wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text_data)

# Display the generated word cloud using Matplotlib
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')  # Turn off the axis
plt.show()
```


Facebook
mentioned
struggling
motivation
day
told
now
great
Im
bed
ðŸ
grandma
go
Hayley
translated
run
lazy
person
auntie
thinks
terrible

```
In [ ]:
```