In [2]: ▶
```python
from textblob import TextBlob
from wordcloud import WordCloud
import pandas as pd
import numpy as np
import re
import matplotlib.pyplot as plt
```

In [90]: ▶
```python
data = pd.read_excel(r"./C_DMK.xlsx",engine='openpyxl')
```

In [91]: ▶
```python
data.head(10)
```

Out[91]:

	Unnamed: 0	user	location	Tweets	stemmed_tweets
0	0	BLawver5	NaN	let slowly walk two made can reread like insan...	['let', 'slowli', 'walk', 'two', 'made', 'can'...
1	1	jcjackson82696	NaN	game give advantage yea way control character ...	['game', 'give', 'advantag', 'yea', 'way', 'co...
2	2	SpadePeoples122	NaN	day strawberry lemonade crack dont yea	['day', 'strawberri', 'lemonad', 'crack', 'don...
3	3	ShmurdaBenny	iAM	one far better thin wearing long even really t...	['one', 'far', 'better', 'thin', 'wear', 'long...
4	4	Everybo96116105	NaN	NaN	[]
5	5	flexyshotjr	NaN	actually might time	['actual', 'might', 'time']
6	6	Everybo96116105	NaN	one sec	['one', 'sec']
7	7	urs_surya09	NaN	let talk	['let', 'talk']
8	8	celine_mnk	Erfurt, Deutschland	vote	['vote']
9	9	Sygic5	NaN	exactly cant win argument bitch go back gold keep	['exact', 'cant', 'win', 'argument', 'bitch', ...

In [92]: ▶
```python
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 496 entries, 0 to 495
Data columns (total 5 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   Unnamed: 0      496 non-null    int64
 1   user            496 non-null    object
 2   location        303 non-null    object
 3   Tweets          464 non-null    object
 4   stemmed_tweets  496 non-null    object
dtypes: int64(1), object(4)
memory usage: 19.5+ KB
```

In [93]: ▶
```python
data['Tweets'] = data['Tweets'].astype('str')
data['stemmed_tweets'] = data['stemmed_tweets'].astype('str')
```

In [94]: ▶
```python
def getSubjectivity(text):
    return TextBlob(text).sentiment.subjectivity
def getPolarity(text):
    return  TextBlob(text).sentiment.polarity
```

In [95]:
```python
data['Subjectivity_st'] = data['stemmed_tweets'].apply(getSubjectivity)
data['Polarity_st'] = data['stemmed_tweets'].apply(getPolarity)
```

In [96]:
```python
data['Subjectivity'] = data['Tweets'].apply(getSubjectivity)
data['Polarity'] = data['Tweets'].apply(getPolarity)
```

In [97]:
```python
df = data
df.head()
```

Out[97]:

	Unnamed: 0	user	location	Tweets	stemmed_tweets	Subjectivity_st	Polarity_st	Subjectivity	Polarity
0	0	BLawver5	NaN	let slowly walk two made can reread like insan...	['let', 'slowli', 'walk', 'two', 'made', 'can'...	0.600000	0.700000	0.650000	0.025000
1	1	jcjackson82696	NaN	game give advantage yea way control character ...	['game', 'give', 'advantag', 'yea', 'way', 'co...	0.644444	0.050000	0.644444	0.050000
2	2	SpadePeoples122	NaN	day strawberry lemonade crack dont yea	['day', 'strawberri', 'lemonad', 'crack', 'don...	0.000000	0.000000	0.000000	0.000000
3	3	ShmurdaBenny	iAM	one far better thin wearing long even really t...	['one', 'far', 'better', 'thin', 'wear', 'long...	0.498810	-0.016667	0.517708	-0.058333
4	4	Everybo96116105	NaN	nan	[]	0.000000	0.000000	0.000000	0.000000

In [98]:
```python
allWords = ' '.join([twts for twts in df['Tweets']])
wordCloud = WordCloud(background_color = "white",width=500, height=300, random_state=21, max_font_size=110).generate(allWords)


plt.imshow(wordCloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```


In [99]:
```python
def getAnalysis(score):
    if score < 0:
        return 'Negative'
    elif score == 0:
        return 'Neutral'
    else:
        return 'Positive'

df['Analysis'] = df['Polarity'].apply(getAnalysis)
```

In [100]: ▶| `df`

Out[100]:

	Unnamed: 0	user	location	Tweets	stemmed_tweets	Subjectivity_st	Polarity_st	Subjectivity
0	0	BLawver5	NaN	let slowly walk two made can reread like insan...	['let', 'slowli', 'walk', 'two', 'made', 'can'...	0.600000	0.700000	0.650000
1	1	jcjackson82696	NaN	game give advantage yea way control character ...	['game', 'give', 'advantag', 'yea', 'way', 'co...	0.644444	0.050000	0.644444
2	2	SpadePeoples122	NaN	day strawberry lemonade crack dont yea	['day', 'strawberri', 'lemonad', 'crack', 'don...	0.000000	0.000000	0.000000
3	3	ShmurdaBenny	iAM	one far better thin wearing long even really t...	['one', 'far', 'better', 'thin', 'wear', 'long...	0.498810	-0.016667	0.517708
4	4	Everybo96116105	NaN	nan	[]	0.000000	0.000000	0.000000
...	...	...	...	...	...	...	...	...
491	491	DylzR6	Hell	probably waffle detect can	['probabl', 'waffl', 'detect', 'can']	0.000000	0.000000	0.000000
492	492	UmarZunnurain	NaN	sammy thank brother already blinded way	['sammi', 'thank', 'brother', 'alreadi', 'blin...	0.666667	-0.500000	0.000000
493	493	mnk_009	ಇಶ್ಟೆ಄ 17.441071,78.437871	b	['b']	0.000000	0.000000	0.000000
494	494	TheOfficalTempz	The skreets of Romford	like	['like']	0.000000	0.000000	0.000000
495	495	ExgIez	NaN	well shouldnt even champ	['well', 'shouldnt', 'even', 'champ']	0.000000	0.000000	0.000000

496 rows × 10 columns

In [101]: ▶|
```
ptweets = df[df.Analysis == 'Positive']
ptweets = ptweets['Tweets']
ptweets

round( (ptweets.shape[0] / df.shape[0]) * 100 , 1)
```

Out[101]: 34.7

In [102]: ▶|
```
allWords = ' '.join([twts for twts in ptweets])
wordCloud = WordCloud(background_color = "white",width=500, height=300, random_state=21, max_font_size=110).generate(allWords)

plt.imshow(wordCloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```

In [103]:
```python
ntweets = df[df.Analysis == 'Negative']
ntweets = ntweets['stemmed_tweets']
ntweets

round( (ntweets.shape[0] / df.shape[0]) * 100, 1)
```

Out[103]: 19.0

In [104]:
```python
ntweets = df[df.Analysis == 'Negative']
ntweets = ntweets['Tweets']
ntweets

round( (ntweets.shape[0] / df.shape[0]) * 100, 1)
```

Out[104]: 19.0

In [105]:
```python
allWords = ' '.join([twts for twts in ntweets])
wordCloud = WordCloud(background_color = "white",max_words = 200,width=500, height=300, random_state=21, max_font_size=110).generate(allWords)

plt.imshow(wordCloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```


In [106]:
```python
neutweets = df[df.Analysis == 'Neutral']
neutweets = neutweets['Tweets']
neutweets

round( (neutweets.shape[0] / df.shape[0]) * 100, 1)
```

Out[106]: 46.4

In [107]: ▶
```python
allWords = ' '.join([twts for twts in neutweets])
wordCloud = WordCloud(background_color = "white",width=500, height=300, random_state=21, max_font_size=110).generate(allWords)


plt.imshow(wordCloud, interpolation="bilinear")
plt.axis('off')
plt.show()
```


In [ ]: ▶