

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import re
import nltk
from nltk.corpus import stopwords
from nltk.stem import WordNetLemmatizer
nltk.download('stopwords')
nltk.download('wordnet')

```

```

[nltk_data] Downloading package stopwords to
[nltk_data] C:\Users\khada\AppData\Roaming\nltk_data...
[nltk_data] Package stopwords is already up-to-date!
[nltk_data] Downloading package wordnet to
[nltk_data] C:\Users\khada\AppData\Roaming\nltk_data...
[nltk_data] Package wordnet is already up-to-date!

```

True

```

from textblob import TextBlob
from tensorflow.keras.preprocessing.text import Tokenizer
from nltk.sentiment.util import *

```

```

trump = pd.read_csv('hashtag_donaldtrump.csv',lineterminator = '\n')
biden = pd.read_csv('hashtag_joebiden.csv',lineterminator = '\n')

```

trump

	created_at	tweet_id	\
0	2020-10-15 00:00:01	1.316529e+18	
1	2020-10-15 00:00:01	1.316529e+18	
2	2020-10-15 00:00:02	1.316529e+18	
3	2020-10-15 00:00:02	1.316529e+18	
4	2020-10-15 00:00:08	1.316529e+18	
...	
725649	2020-11-05 23:59:59	1.324502e+18	
725650	2020-11-05 23:59:59	1.324502e+18	
725651	2020-11-05 23:59:59	1.324502e+18	
725652	2020-11-05 23:59:59	1.324502e+18	
725653	2020-11-05 23:59:59	1.324502e+18	

	tweet	likes	\
0	#Elecciones2020 En #Florida: #JoeBiden dice ...	0.0	
1	Usa 2020, Trump contro Facebook e Twitter: cop...	26.0	
2	#Trump: As a student I used to hear for years,...	2.0	
3	2 hours since last tweet from #Trump! Maybe he...	0.0	
4	You get a tie! And you get a tie! #Trump 's ra...	4.0	
...	
725649	@DavidCornDC #coup by #trump 🇺🇸	2.0	
725650	Did you ever speak to someone on public transp...	3.0	

725651	So they accepted ballots WAY MORE than three d...	3.0
725652	#Trump zegt nu gewoon ... 'if you only count m...	6.0
725653	#Trump speech truth-o-meter as we speak...\n\n...	1.0

	retweet_count	source	user_id	\
0	0.0	TweetDeck	3.606665e+08	
1	9.0	Social Mediaset	3.316176e+08	
2	1.0	Twitter Web App	8.436472e+06	
3	0.0	Trumptytweeter	8.283556e+17	
4	3.0	Twitter for iPhone	4.741380e+07	
...	
725649	0.0	Twitter for iPhone	1.230104e+18	
725650	2.0	Twitter for Android	2.964605e+07	
725651	0.0	Twitter Web App	3.753528e+08	
725652	0.0	Twitter Web App	2.902307e+08	
725653	0.0	Twitter for Android	2.697767e+09	

	user_name	user_screen_name	\
0	El Sol Latino News	elsollatinonews	
1	Tgcom24	MediasetTgcom24	
2	snarke	snarke	
3	Trumptytweeter	trumptytweeter	
4	Rana Abtar - رنا أبتار	Ranaabtar	
...	
725649	Debra Weller Olson - "VOTE"	OlsonWeller	
725650	Stephen Connolly	5Konnolly87	
725651	Brian Kennedy	GrowingUpHockey	
725652	Tom Du Bois Vanhoutte (B) (E)	Du_Bois	
725653	Robbie "#PR17 forever"	Rob2K	

	user_description	...	\
0	🌐 Noticias de interés para latinos de la costa...	...	
1	Profilo ufficiale di Tgcom24: tutte le notizie...	...	
2	Will mock for food! Freelance writer, blogger,...	...	
3	If he doesn't tweet for some time, should we b...	...	
4	Washington Correspondent, Lebanese-American ,c...	...	
...	
725649	2012-14 wkd 3Blocks "S"of #WallSt NYC "GC"Firm...	...	
725650		NaN	...
725651	Montreal-born, live in Cali. I wrote Growing U...	...	
725652	RT&like&endorse "#Nationalism eq. #poison, yo...	...	
725653	@TheBharatArmy @bccci @CTFCOfficial fan #COYR @...	...	

	user_followers_count	user_location	lat	\
0	1860.0	Philadelphia, PA / Miami, FL	25.774270	
1	1067661.0	NaN	NaN	

2	1185.0	Portland	45.520247
3	32.0	NaN	NaN
4	5393.0	Washington DC	38.894992
...
725649	5654.0	United States	39.783730
725650	539.0	Drumlane / Portmarnock	NaN
725651	1092.0	Canadian-Californian	NaN
725652	2952.0	@Nottinghill @Antwerp	NaN
725653	600.0	London, England	51.507322

continent \	long	city	country	
0 America	-80.193660	NaN	United States of America	North
1 NaN	NaN	NaN	NaN	
2 America	-122.674195	Portland	United States of America	North
3 NaN	NaN	NaN	NaN	
4 America	-77.036558	Washington	United States of America	North
...
725649 America	-100.445882	NaN	United States	North
725650 NaN	NaN	NaN	NaN	
725651 NaN	NaN	NaN	NaN	
725652 NaN	NaN	NaN	NaN	
725653 Europe	-0.127647	London	United Kingdom	

	state	state_code	collected_at
0	Florida	FL	2020-10-21 00:00:00
1	NaN	NaN	2020-10-21 00:00:00.373216530

2	Oregon	OR	2020-10-21 00:00:00.746433060
3	NaN	NaN	2020-10-21 00:00:01.119649591
4	District of Columbia	DC	2020-10-21 00:00:01.492866121
...
725649	NaN	NaN	2020-11-06 09:50:53.176351
725650	NaN	NaN	2020-11-06 09:50:53.272491
725651	NaN	NaN	2020-11-06 09:50:53.108048
725652	NaN	NaN	2020-11-06 09:50:53.401021
725653	England	ENG	2020-11-06 09:50:53.372230

























[725654 rows x 21 columns]

biden

	created_at	tweet_id \
0	2020-10-15 00:00:01	1.316529e+18
1	2020-10-15 00:00:18	1.316529e+18
2	2020-10-15 00:00:20	1.316529e+18
3	2020-10-15 00:00:21	1.316529e+18
4	2020-10-15 00:00:22	1.316529e+18
...
499823	2020-11-05 23:59:45	1.324502e+18
499824	2020-11-05 23:59:47	1.324502e+18
499825	2020-11-05 23:59:50	1.324502e+18
499826	2020-11-05 23:59:54	1.324502e+18
499827	2020-11-05 23:59:58	1.324502e+18

	tweet	likes \
0	#Elecciones2020 En #Florida: #JoeBiden dice ...	0.0
1	#HunterBiden #HunterBidenEmails #JoeBiden #Joe...	0.0
2	@IslandGirlPRV @BradBeauregardJ @MeidasTouch T...	0.0
3	@chrislongview Watching and setting dvr. Let's...	0.0
4	#censorship #HunterBiden #Biden #BidenEmails #...	1.0
...
499823	Votes get "whittled down"? Counting of mail i...	0.0
499824	CONTINUAN SALIENDO LAS PRUEBAS DEL FRAUDE ELEC...	1.0
499825	#ElectionResults2020 #USElection #ThursdayThou...	1.0
499826	I wonder who they're voting for 🤔... - #fyp #f...	0.0
499827	Biden - 290 \nTrump 🤔 - 248 \nEasy money babby...	0.0

	retweet_count	source	user_id \
0	0.0	TweetDeck	3.606665e+08
1	0.0	Twitter for iPad	8.099044e+08
2	0.0	Twitter Web App	3.494182e+09
3	0.0	Twitter for iPhone	8.242596e+17
4	0.0	Twitter Web App	1.032807e+18
...
499823	0.0	Twitter for iPad	8.059161e+08
499824	1.0	Twitter for iPhone	4.805877e+08
499825	0.0	Twitter for iPhone	2.265816e+09
499826	0.0	Instagram	8.317075e+17
499827	0.0	Twitter for iPhone	9.071075e+17

	user_name	user_screen_name	\
0	El Sol Latino News	elsollatinonews	
1	Cheri A. 	Biloximeemaw	
2	Flag Waver	Flag_Wavers	
3	Michelle Ferg	MichelleFerg4	
4	the Gold State	thegoldstate	
...			
499823	Sherry Slater	NS_Slater	
499824	Portero Mexico	PorteroMexico	
499825	VOTE EARLY!                       		

```

                                user_description    ... \
0      🌐 Noticias de interés para latinos de la costa...   ...
1      Locked and loaded Meemaw. Love God, my family     ...
2                                                         NaN    ...
3                                                         NaN    ...
4      A Silicon Valley #independent #News #Media #St...   ...
...                                                     ...    ...
499823 Mom of 4 great kids - hard working, proud Demo...   ...
499824                                                         NaN    ...
499825 If my content resonates—follow me. Resist •Pro...   ...
499826 ♂ : IG 📺 NathanKTV \n○: ➡️👉🏻https://t.co/o...   ...
499827 crypto degenerate : truth seeker : datsun enth...   ...

```

	user_followers_count	user_location	
0	1860.0	Philadelphia, PA / Miami, FL	25.774270
1	6628.0	NaN	NaN
2	1536.0	Golden Valley Arizona	46.304036

3	27.0	NaN	NaN
4	390.0	California, USA	36.701463
...
499823	27.0	NaN	NaN
499824	873.0	NaN	NaN
499825	140.0	Earth	NaN
499826	0.0	Beverly Hills, CA	34.069650
499827	94.0	San Francisco, CA	37.779026

continent \	long	city	country
0 America	-80.193660	NaN	United States of America
1 NaN	NaN	NaN	NaN
2 America	-109.171431	NaN	United States of America
3 NaN	NaN	NaN	NaN
4 America	-118.755997	NaN	United States of America
...
...
499823 NaN	NaN	NaN	NaN
499824 NaN	NaN	NaN	NaN
499825 NaN	NaN	NaN	NaN
499826 America	-118.396306	NaN	United States of America
499827 America	-122.419906	San Francisco	United States of America

	state	state_code	collected_at
0	Florida	FL	2020-10-21 00:00:00
1	NaN	NaN	2020-10-21 00:00:00.517827283
2	Montana	MT	2020-10-21 00:00:01.035654566
3	NaN	NaN	2020-10-21 00:00:01.553481849
4	California	CA	2020-10-21 00:00:02.071309132
...

```

499823      NaN      NaN      2020-11-06 10:04:37.074759
499824      NaN      NaN      2020-11-06 10:04:36.811037
499825      NaN      NaN      2020-11-06 10:04:37.040510
499826  California      CA      2020-11-06 10:04:36.975515
499827  California      CA      2020-11-06 10:04:37.088501

```

```
[499828 rows x 21 columns]
```

```

print(trump.shape)
print(biden.shape)

```

```

(725654, 21)
(499828, 21)

```

```
trump.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 725654 entries, 0 to 725653
Data columns (total 21 columns):

```

#	Column	Non-Null Count	Dtype
0	created_at	725654 non-null	object
1	tweet_id	725654 non-null	float64
2	tweet	725654 non-null	object
3	likes	725654 non-null	float64
4	retweet_count	725654 non-null	float64
5	source	724897 non-null	object
6	user_id	725654 non-null	float64
7	user_name	725639 non-null	object
8	user_screen_name	725654 non-null	object
9	user_description	651481 non-null	object
10	user_join_date	725654 non-null	object
11	user_followers_count	725654 non-null	float64
12	user_location	505267 non-null	object
13	lat	342071 non-null	float64
14	long	342071 non-null	float64
15	city	172431 non-null	object
16	country	339877 non-null	object
17	continent	339890 non-null	object
18	state	248088 non-null	object
19	state_code	233393 non-null	object
20	collected_at	725654 non-null	object

```
dtypes: float64(7), object(14)
```

```
memory usage: 116.3+ MB
```

```
trump['candidate'] = 'Trump'
```

```
biden['candidate'] = 'Biden'
```

```
df = pd.concat([trump,biden])
```

```
df
```


499826	NathanKTV	NathanKTV_
499827	BraddahBit	alohacory_

```

                                user_description    ... \
0      🌐 Noticias de interés para latinos de la costa...   ...
1      Profilo ufficiale di Tgcom24: tutte le notizie...   ...
2      Will mock for food! Freelance writer, blogger,...   ...
3      If he doesn't tweet for some time, should we b...   ...
4      Washington Correspondent, Lebanese-American ,c...   ...
...                                     ...               ...
499823 Mom of 4 great kids - hard working, proud Demo...   ...
499824                                         NaN           ...
499825 If my content resonates—follow me. Resist •Pro...   ...
499826 ♂ : IG 📺 NathankTV \n○: ➡️👉🔗https://t.co/o...   ...
499827 crypto degenerate : truth seeker : datsun enth...   ...

```

city \	user_location	lat	long
0	Philadelphia, PA / Miami, FL	25.774270	-80.193660
NaN			
1	NaN	NaN	NaN
NaN			
2	Portland	45.520247	-122.674195
Portland			
3	NaN	NaN	NaN
NaN			
4	Washington DC	38.894992	-77.036558
Washington			
...
...			
499823	NaN	NaN	NaN
NaN			
499824	NaN	NaN	NaN
NaN			
499825	Earth	NaN	NaN
NaN			
499826	Beverly Hills, CA	34.069650	-118.396306
NaN			
499827	San Francisco, CA	37.779026	-122.419906
San Francisco			San Francisco

	country	continent	state
0	United States of America	North America	Florida
1	NaN	NaN	NaN
2	United States of America	North America	Oregon
3	NaN	NaN	NaN

4	United States of America	North America	District of Columbia
...			
499823		NaN	NaN
499824		NaN	NaN
499825		NaN	NaN
499826	United States of America	North America	California
499827	United States of America	North America	California

	state_code	collected_at	candidate
0	FL	2020-10-21 00:00:00	Trump
1	NaN	2020-10-21 00:00:00.373216530	Trump
2	OR	2020-10-21 00:00:00.746433060	Trump
3	NaN	2020-10-21 00:00:01.119649591	Trump
4	DC	2020-10-21 00:00:01.492866121	Trump
...
499823	NaN	2020-11-06 10:04:37.074759	Biden
499824	NaN	2020-11-06 10:04:36.811037	Biden
499825	NaN	2020-11-06 10:04:37.040510	Biden
499826	CA	2020-11-06 10:04:36.975515	Biden
499827	CA	2020-11-06 10:04:37.088501	Biden

[1225482 rows x 22 columns]

df.shape

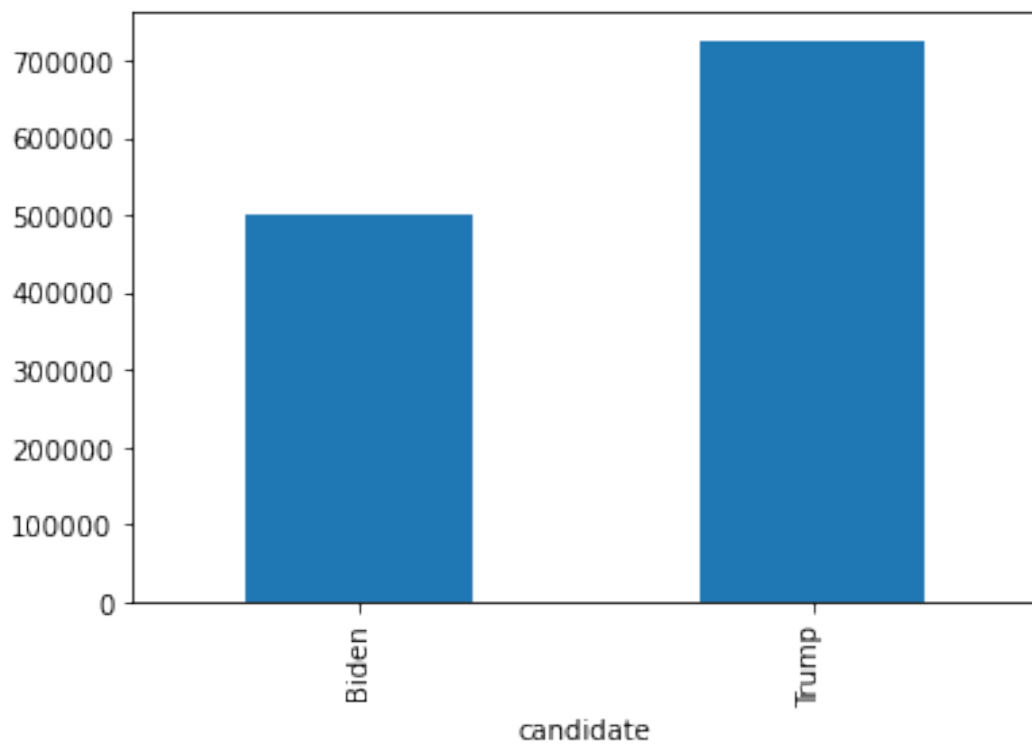
(1225482, 22)

EDA

in visual who have got more number of tweets

df.groupby('candidate')['tweet'].count().plot.bar()

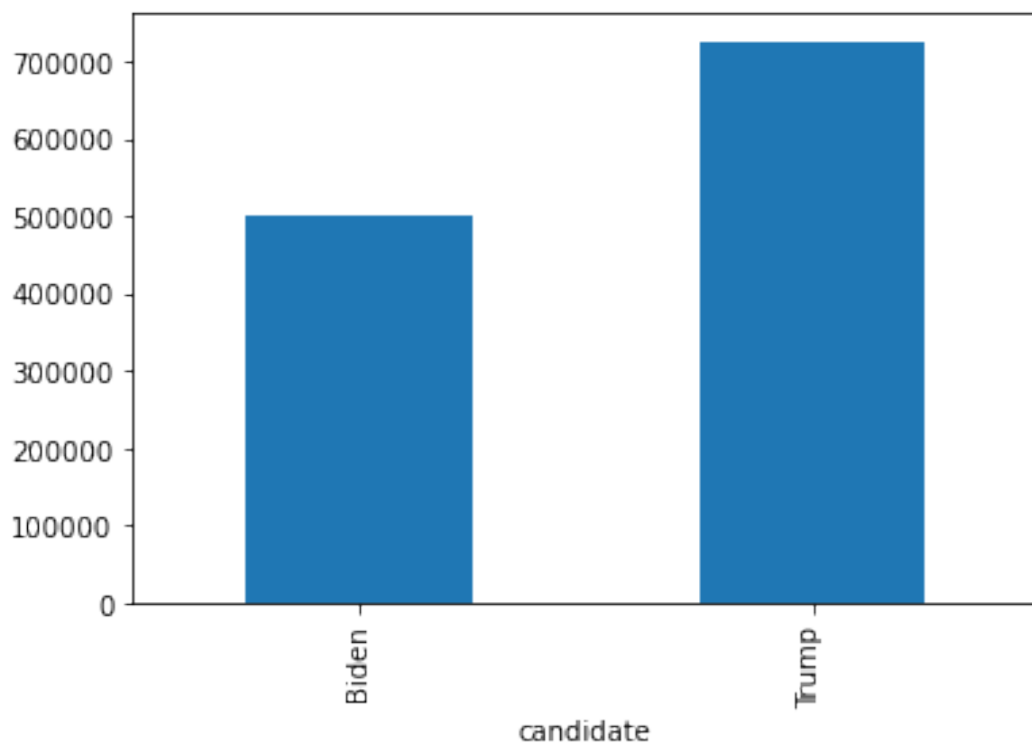
<matplotlib.axes._subplots.AxesSubplot at 0x1d483449610>



comparison of likes

```
df.groupby('candidate')['likes'].count().plot.bar()
```

<matplotlib.axes._subplots.AxesSubplot at 0x1d485515790>



```
df.country.value_counts()

United States of America    262250
United States               50846
United Kingdom              35946
Germany                    23590
France                     22459
...
Benin                       2
Turks and Caicos Islands    1
Antigua and Barbuda         1
Saint Lucia                 1
Northern Mariana Islands     1
Name: country, Length: 185, dtype: int64
```

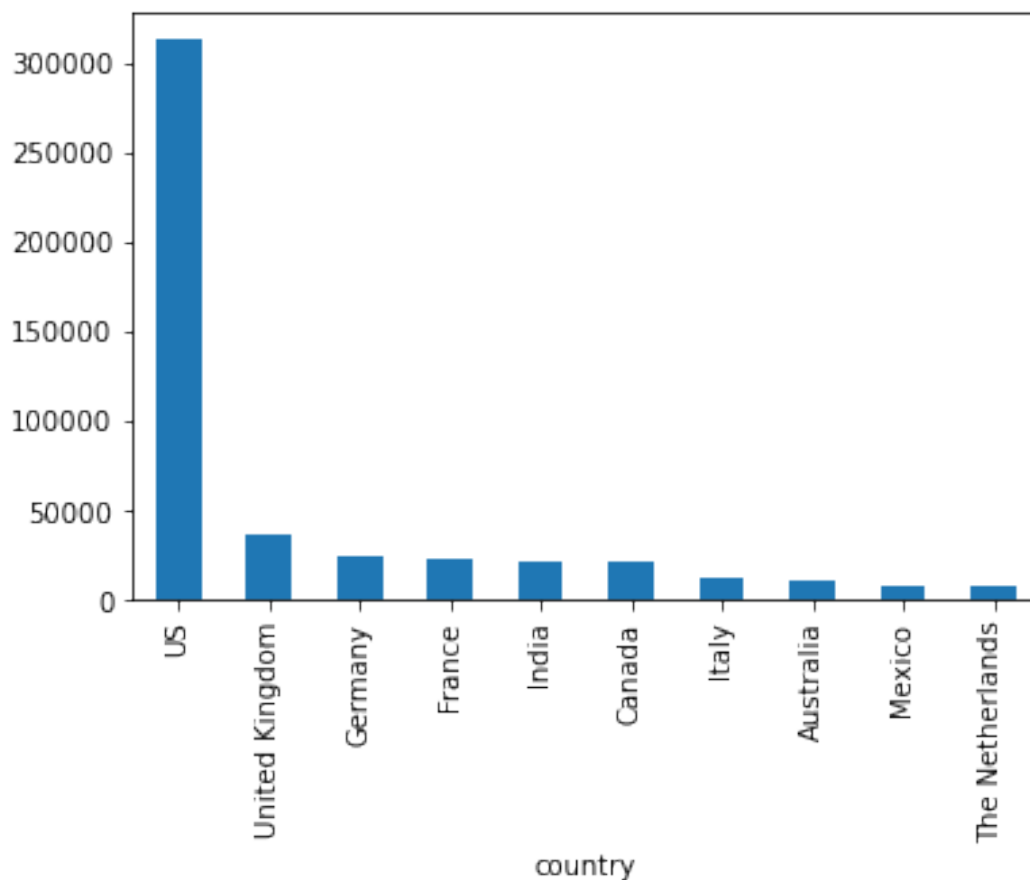
```
df['country'] = df['country'].replace({'United States of America': 'US', 'United States': 'US'})
```

```
df.country.value_counts()

US                313096
United Kingdom    35946
Germany           23590
France            22459
India             21891
...
South Sudan       2
Antigua and Barbuda 1
Saint Lucia       1
Northern Mariana Islands 1
Turks and Caicos Islands 1
Name: country, Length: 184, dtype: int64
```

```
# top 10 countries with highest number of tweets - bar graph
df.groupby('country')
['tweet'].count().sort_values(ascending=False).head(10).plot.bar()

<matplotlib.axes._subplots.AxesSubplot at 0x1d485599eb0>
```



```
# tweets for biden and trump from each country
top10country = df.groupby('country')
['tweet'].count().sort_values(ascending=False).head(10).index.tolist()
top10country

['US',
 'United Kingdom',
 'Germany',
 'France',
 'India',
 'Canada',
 'Italy',
 'Australia',
 'Mexico',
 'The Netherlands']

tweet_df = df.groupby(['country', 'candidate'])
['tweet'].count().reset_index()
tweet_df
```

	country	candidate	tweet
0	Afghanistan	Biden	53
1	Afghanistan	Trump	78
2	Albania	Biden	29

3	Albania	Trump	55
4	Algeria	Biden	9
..
354	Yemen	Trump	78
355	Zambia	Biden	8
356	Zambia	Trump	26
357	Zimbabwe	Biden	33
358	Zimbabwe	Trump	43

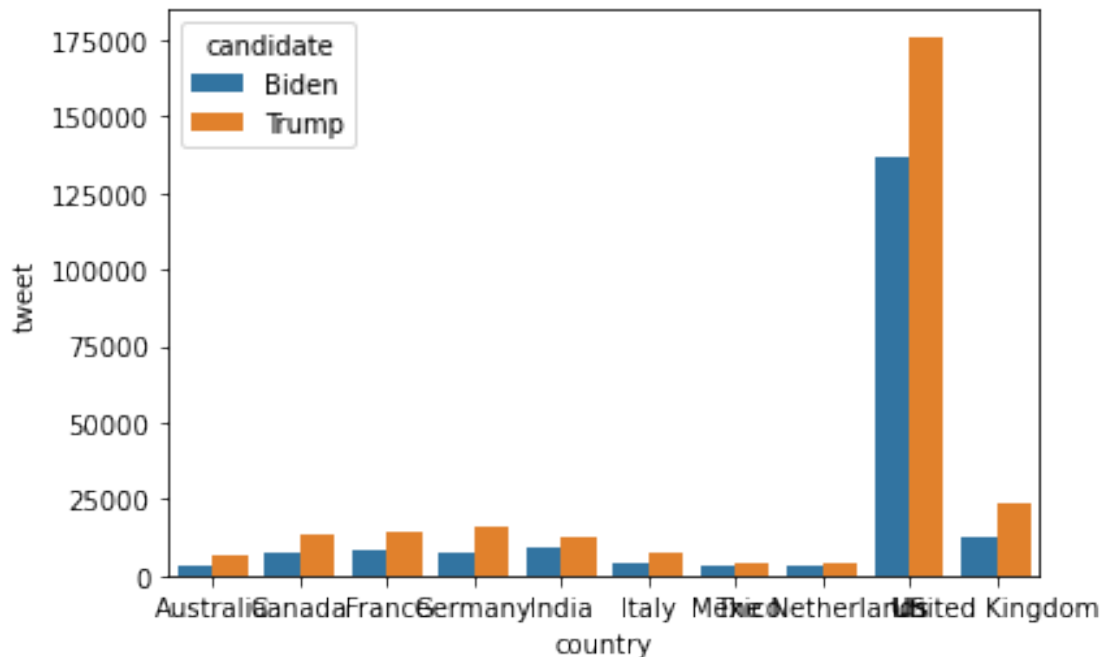
[359 rows x 3 columns]

```
tweet_df = tweet_df[tweet_df['country'].isin(top10country)]
tweet_df
```

	country	candidate	tweet
15	Australia	Biden	3601
16	Australia	Trump	6691
52	Canada	Biden	7401
53	Canada	Trump	13320
109	France	Biden	8244
110	France	Trump	14215
115	Germany	Biden	7503
116	Germany	Trump	16087
143	India	Biden	9155
144	India	Trump	12736
157	Italy	Biden	4442
158	Italy	Trump	7316
202	Mexico	Biden	2930
203	Mexico	Trump	4237
321	The Netherlands	Biden	2812
322	The Netherlands	Trump	4210
331	US	Biden	136920
332	US	Trump	176176
339	United Kingdom	Biden	12626
340	United Kingdom	Trump	23320

```
sns.barplot(data=tweet_df,x='country',y='tweet',hue='candidate')
```

<matplotlib.axes._subplots.AxesSubplot at 0x1d48560cd90>



```
lm= WordNetLemmatizer()
```

```
def clean(text):
    text = str(text).lower()
    text = re.sub('[^a-z]', ' ', str(text))
    text = text.split()
    text = [lm.lemmatize(j) for j in text if j not in
set(stopwords.words('english'))]
    text = ' '.join(str(k) for k in text)
    return text
```

```
# sentiment analysis
```

```
trump_tweets =
df.query('(candidate=="Trump")').sort_values('user_followers_count',as
cending=False).drop_duplicates(['user_id'])[['tweet','country']]
```

```
trump_tweets = trump_tweets.loc[trump_tweets.country=='US']
```

```
trump_tweets.isnull().sum()
```

```
tweet      0
country    0
dtype: int64
```

```
trump_tweets['cleantext'] = trump_tweets['tweet'].apply(clean)
```

```
trump_tweets
```

	tweet	country	\
115369	President Donald #Trump's youngest daughter, T...	US	
270220	Woah. Have you read this article?\n\nAll the p...	US	
661169	#SethMeyers lambastes Fox News for parroting #...	US	
693531	✧ Las #noticias más importantes de la mañana: ...	US	
566989	#Trump Habla en segundos. Todavía nada está de...	US	
...	
205989	About the #fracking comment last night. Is Jo...	US	
462011	Dirty Don trying to cheat at the polls. @realD...	US	
68033	The United States shows more DEATHS than other...	US	
228347	#iamvalued #youarevaued #msc #Trump #trump2020...	US	
205868	Picture from the last night #Debates2020 #Bide...	US	

	cleantext
115369	president donald trump youngest daughter tiffa...
270220	woah read article president debt donald trump ...
661169	sethmeyers lambastes fox news parroting donald...
693531	la noticias importantes de la ana sigue el con...
566989	trump habla en segundos todav nada est definid...
...	...
205989	fracking comment last night joe biden bidenhar...
462011	dirty trying cheat poll realdonaldtrump donald...
68033	united state show death country lamestream fak...
228347	iamvalued youarevaued msc trump trump seat sti...
205868	picture last night debate biden trump http co ...

[39630 rows x 3 columns]

Create function for polarity and subjectivity

Subjectivity - if the sentence is subjective or objective - subjective refers to the personal opinions, emotions or peoples's judgement wheras objectrive refers to factual information.

Polarity - float value lies between[-1,1] where 1 means positive and -1 means negative

```
def getsubjectivity(text):
    return TextBlob(text).sentiment.subjectivity
def getpolarity(text):
    return TextBlob(text).sentiment.polarity
def getAnalysis(score):
    if score<0:
        return 'negative'
    elif score==0:
        return 'neutral'
    else:
        return 'positive'
```

```
trump_tweets['subjectivity'] =
trump_tweets['cleantext'].apply(getsubjectivity)
```



```
trump_tweets['polarity'] =
trump_tweets['cleantext'].apply(getpolarity)
```

```
trump_tweets['sentiment'] =
trump_tweets['polarity'].apply(getAnalysis)
```

```
trump_tweets
```

	tweet	country	\
115369	President Donald #Trump's youngest daughter, T...	US	
270220	Woah. Have you read this article?\n\nAll the p...	US	
661169	#SethMeyers lambastes Fox News for parroting #...	US	
693531	✂ Las #noticias más importantes de la mañana: ...	US	
566989	#Trump Habla en segundos. Todavía nada está de...	US	
...	
205989	About the #fracking comment last night. Is Jo...	US	
462011	Dirty Don trying to cheat at the polls. @realD...	US	
68033	The United States shows more DEATHS than other...	US	
228347	#iamvalued #youarevaued #msc #Trump #trump2020...	US	
205868	Picture from the last night #Debates2020 #Bide...	US	

```
clean text
```

```
subjectivity \
```

115369	president donald trump youngest daughter tiffa...	0.000000
270220	woah read article president debt donald trump ...	0.000000
661169	sethmeyers lambastes fox news parroting donald...	0.000000
693531	la noticias importantes de la ana sigue el con...	0.000000
566989	trump habla en segundos todav nada est definid...	0.000000

```
... ..
```

205989	fracking comment last night joe biden bidenhar...	0.066667
462011	dirty trying cheat poll realdonaldtrump donald...	0.800000
68033	united state show death country lamestream fak...	0.733333
228347	iamvalued youarevaued msc trump trump seat sti...	0.000000
205868	picture last night debate biden trump http co ...	0.066667

	polarity	sentiment
115369	0.00	neutral
270220	0.00	neutral
661169	0.00	neutral
693531	0.00	neutral

```

566989      0.00  neutral
...
205989      0.00  neutral
462011     -0.60  negative
68033      -0.44  negative
228347      0.00  neutral
205868      0.00  neutral

```

```
[39630 rows x 6 columns]
```

```

biden_tweets =
df.query('(candidate=="Biden")').sort_values('user_followers_count',as
cending=False).drop_duplicates(['user_id'])[['tweet','country']]
biden_tweets = biden_tweets.loc[biden_tweets.country=='US']

```

```
biden_tweets['cleantext'] = biden_tweets['tweet'].apply(clean)
```

```

biden_tweets['polarity'] =
biden_tweets['cleantext'].apply(getpolarity)

```

```

biden_tweets['subjectivity'] =
biden_tweets['cleantext'].apply(getsubjectivity)

```

```

biden_tweets['sentiment'] =
biden_tweets['polarity'].apply(getAnalysis)

```

```
biden_tweets
```

	tweet	country	\
175650	#JoeBiden is enlisting A-list star power to he...	US	
343145	The big fight TODAY!!!! @realDonaldTrump vs #...	US	
459947	#JimmyKimmel offered his view about how the la...	US	
493728	Sigue el conteo de votos en #Georgia: #Trump y...	US	
281536	Are you there, God? It's me again, Bette. We'r...	US	
...	
484098	UNITED WE STAND PEOPLE don't let the clowns ru...	US	
67722	@TheJeffBridges And, I voted for the only hono...	US	
313770	#2020ElectionPrediction\n\n#Trump wins 300-320...	US	
42480	The world is sick and needs a hero. I think #b...	US	
427600	This race is still tight!! Biden has a tiny le...	US	

	cleantext	polarity	\
175650	joe Biden enlisting list star power help close ...	0.000000	
343145	big fight today realdonaldtrump v joe Biden got...	0.000000	
459947	jimmykimmel offered view last day gone close r...	0.000000	
493728	sigue el conteo de votos en georgia trump bide...	0.000000	
281536	god bette scared fascist whitehouse racist cri...	0.100000	
...	
484098	united stand people let clown running presiden...	0.083333	
67722	thejeffbridges voted honorable man running pre...	0.000000	
313770	electionprediction trump win electoral vote ma...	0.733333	

```
42480    world sick need hero think biden way towards s... -0.714286
427600    race still tight biden tiny lead michigan neva... 0.255357
```

```

      subjectivity sentiment
175650      0.000000    neutral
343145      0.100000    neutral
459947      0.066667    neutral
493728      0.000000    neutral
281536      0.562500    positive
...
484098      0.541667    positive
67722      0.000000    neutral
313770      0.566667    positive
42480      0.857143    negative
427600      0.446429    positive

```

```
[42976 rows x 6 columns]
```

```
trump_tweets.sentiment.value_counts(normalize=True)*100
```

```
neutral      44.789301
positive     34.625284
negative     20.585415
Name: sentiment, dtype: float64
```

```
biden_tweets.sentiment.value_counts(normalize=True)*100
```

```
neutral      46.854058
positive     36.827532
negative     16.318410
Name: sentiment, dtype: float64
```

```

from wordcloud import WordCloud, STOPWORDS
def word_cloud(wd_list):
    stopwords = set(STOPWORDS)
    all_words = ' '.join([text for text in wd_list])
    wordcloud = WordCloud(background_color='white',stopwords =
stopwords,
                        width=1600,height=800,
max_font_size=800,max_words=100).generate(all_words)
    plt.figure(figsize=(12,10))
    plt.axis('off')
    plt.imshow(wordcloud)

```

```
word_cloud(trump_tweets['tweet'][:5000])
```

