

```
In [2]:
             import re
             import nltk
             import string
             import warnings
             import numpy as np
             import pandas as pd
             import seaborn as sns
             import matplotlib.pyplot as plt
            pd.set option("display.max colwidth", 200)
            warnings.filterwarnings("ignore", category = DeprecationWarning)
             %matplotlib inline
In [3]:
            train = pd.read csv("train tweets.csv")
            test = pd.read csv("test tweets.csv")
In [4]:
             train
Out [4]:
                         id label
                                                                                                       tweet
                                      Quser when a father is dysfunctional and is so selfish he drags his kids
                 0
                          1
                                 0
                                                                                   into his dysfunction. #run
                                        @user @user thanks for #lyft credit i can't use cause they don't offer
                          2
                                 0
                 1
                                                          wheelchair vans in pdx. #disapointed #getthanked
                 2
                          3
                                 0
                                                                                         bihday your majesty
                                        #model i love u take with u all the time in urðŸ"±!!! ðŸ~™ðŸ~ŽðŸ'"ðŸ'...
                                 0
                 3
                                                                                                  ðΫʹ¦ðΫʹ¦ðΫʹ¦
                          5
                                 0
                                                                         factsguide: society now #motivation
                                                                                   ate @user isz that youuu?
            31957 31958
                                 0
                                                  \eth\ddot{Y}^{\sim}\quad \eth\ddot{Y}^{\sim}\quad \eth\ddot{Y}^{\sim}\quad \eth\ddot{Y}^{\sim}\quad \eth\ddot{Y}^{\sim}\quad \eth\ddot{Y}^{\sim}\quad \eth\ddot{Y}^{\sim}\quad \eth\ddot{Y}^{\sim}\quad \mathring{a}\quad \text{mi,}
                                       to see nina turner on the airwaves trying to wrap herself in the mantle
            31958 31959
                                 0
                                                   of a genuine hero like shirley chisolm. #shame #imwithher
            31959 31960
                                 0
                                              listening to sad songs on a monday morning otw to work is sad
            31960 31961
                                 1
                                         @user #sikh #temple vandalised in in #calgary, #wso condemns act
            31961 31962
                                 0
                                                                              thank you @user for you follow
          31962 rows × 3 columns
In [5]:
             test
Out[5]:
                         id
                                                                                                       tweet
```

0	31963	My-personal-files/twitter_sentiment Analysis.ipynb at main · saipranay190198/My-personal-files #studione #aisine #requires #passion #dedication #winpower to nind #newmaterials…
1	31964	@user #white #supremacists want everyone to see the new â€~ #birds' #movie â€" and here's why
2	31965	safe ways to heal your #acne!! #altwaystoheal #healthy #healing!!
3	31966	is the hp and the cursed child book up for reservations already? if yes, where? if no, when? ðŸ~ ðŸ~ ðŸ~ #harrypotter #pottermore #favorite
4	31967	3rd #bihday to my amazing, hilarious #nephew eli ahmir! uncle dave loves you and misses…
17192	49155	thought factory: left-right polarisation! #trump #uselections2016 #leadership #politics #brexit #blm >3
17193	49156	feeling like a mermaid ðŸ~~ #hairflip #neverready #formal #wedding #gown #dresses #mermaid …

#hillary #campaigned today in #ohio((omg)) & amp; used words like

my song "so glad" free download! #shoegaze #newmusic #newsong

organizations #work #mindset

"assets&liability" never once did #clinton say thee(word) #radicalization

happy, at work conference: right mindset leads to culture-of-development

17197 rows × 2 columns

49157

17195 49158

17196 49159

17194

```
In [6]:
         train[train['label'] == 0].head(10)
```

Out[6]:	id label		label	tweet			
	0	1	0	@user when a father is dysfunctional and is so selfish he drags his kids into his dysfunction. #run			
	1	2	0	@user @user thanks for #lyft credit i can't use cause they don't offer wheelchair vans in pdx. #disapointed #getthanked			
	2 3 0		0	bihday your majesty			
;		4	0	#model i love u take with u all the time in urðŸ"±!!! ðŸ~™ðŸ~ŽðŸ'"ðŸ'ðŸ'¦ðŸ'¦			
	4	5	0	factsguide: society now #motivation			
5 6		0	[2/2] huge fan fare and big talking before they leave. chaos and pay disputes when they get there. #allshowandnogo				
	6	7	0	@user camping tomorrow @user @user @user @user @user @user danny…			
	7	8	0	the next school year is the year for exams.ðŸ~¯ can't think about that ðŸ~− #school #exams #hate #imagine #actorslife #revolutionschool #girl			
	8	9	0	we won!!! love the land!!! #allin #cavs #champions #cleveland #clevelandcavaliers …			
	9	10	0	@user @user welcome here!i'm it's so #gr8!			

```
In [7]:
         train[train['label'] == 1].head(10)
```

```
label
 Out[7]:
                 id
                                                                                            tweet
                                  @user #cnn calls #michigan middle school 'build the wall' chant " #tcot
            13
                 14
                         1
                               no comment! in #australia #opkillingbay #seashepherd #helpcovedolphins
                         1
            14
                 15
                                                                        #thecove #helpcovedolphins
            17
                 18
                         1
                                                                               retweet if you agree!
                 24
                         1
            23
                                                       @user @user lumpy says i am a . prove it lumpy.
                            it's unbelievable that in the 21st century we'd need something like this. again.
                 35
                         1
            34
                                                                           #neverump #xenophobia
                 57
                         1
                                                                @user lets fight against #love #peace
            56
                                   õŸ~©the white establishment can't have blk folx running around loving
                         1
            68
                 69
                                                             themselves and promoting our greatness
                            @user hey, white people: you can call people 'white' by @user #race #identity
                         1
                 78
            77
                                                                                          #medâ€!
                 83
                         1
                               how the #altright uses & to lure men into #whitesupremacy
            82
                             @user i'm not interested in a #linguistics that doesn't address #race & .
            111
                112
                         1
                                                    racism is about #power. #raciolinguistics brings…
 In [8]:
            train.shape , test.shape
           ((31962, 3), (17197, 2))
Out[8]:
 In [9]:
            train["label"].value counts()
                 29720
Out[9]:
                   2242
           Name: label, dtype: int64
In [10]:
            length train = train["tweet"].str.len()
            length test = test["tweet"].str.len()
            plt.hist(length train, bins = 20, label = "train tweets")
            plt.hist(length test, bins = 20, label = "test tweets")
            plt.legend()
            plt.show()
            6000
                                                          train tweets
                                                           test tweets
            5000
            4000
            3000
            2000
            1000
```

```
9/12/22, 9:21 PM
```

```
In [11]:
            combine = train.append(test, ignore index = True)
            combine.shape
           C:\Users\rsaipranayreddy\AppData\Local\Temp\ipykernel_12188\3655217207.p
           y:1: FutureWarning: The frame.append method is deprecated and will be rem
           oved from pandas in a future version. Use pandas.concat instead.
             combine = train.append(test, ignore index = True)
           (49159, 3)
Out[11]:
In [12]:
            def remove_pattern(input_txt, pattern):
                r = re.findall(pattern, input txt)
                for i in r:
                     input_txt = re.sub(i,'',input_txt)
                     return input_txt
                else:
                     return input txt
In [13]:
            combine['tidy_tweet'] = np.vectorize(remove_pattern)(combine['tweet'],
            combine.head()
              id label
                                                       tweet
                                                                                    tidy_tweet
Out[13]:
                         @user when a father is dysfunctional and
                                                              when a father is dysfunctional and is
               1
                   0.0
                            is so selfish he drags his kids into his
                                                                so selfish he drags his kids into his
                                             dysfunction. #run
                                                                               dysfunction. #run
                             @user @user thanks for #lyft credit i
                                                                  thanks for #lyft credit i can't use
                                 can't use cause they don't offer
                                                                 cause they don't offer wheelchair
               2
                    0.0
                            wheelchair vans in pdx. #disapointed
                                                                        vans in pdx. #disapointed
                                                 #getthanked
                                                                                   #getthanked
           2
               3
                    0.0
                                           bihday your majesty
                                                                             bihday your majesty
                                                                 #model i love u take with u all the
                         #model i love u take with u all the time in
                                                                time in urðŸ"±!!! ðŸ~™ðŸ~ŽðŸ'"ðŸ'...
                    0.0
                           urðŸ"±!!! ðŸ~™ðŸ~ŽðŸ'"ðŸ'...ðŸ'¦ðŸ'¦ðŸ'¦
                                                                                     ðΫʹ¦ðΫʹ¦ðΫʹ¦
               5
                    0.0
                             factsguide: society now #motivation
                                                               factsguide: society now #motivation
In [14]:
            combine['tidy tweet'] = combine['tidy tweet'].str.replace("[^a-zA-Z#]",
            combine.head(10)
           C:\Users\rsaipranayreddy\AppData\Local\Temp\ipykernel 12188\1044572875.p
           y:1: FutureWarning: The default value of regex will change from True to F
           alse in a future version.
             combine['tidy tweet'] = combine['tidy tweet'].str.replace("[^a-zA-Z#]",
Out[14]:
              id label
                                                      tweet
                                                                                    tidy_tweet
                            @user when a father is dysfunctional
                                                              when a father is dysfunctional and is
           0
               1
                    0.0
                           and is so selfish he drags his kids into
                                                                so selfish he drags his kids into his
                                         his dysfunction. #run
                                                                               dysfunction #run
```

ut[15]: n [16]: ut[16]:	0 1 2 3 4	1 2 3 4 5 oken	ized_	@user when a father is dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and waste for the disapoint of the cause they don't offer wheelchair vans in #disapointed #getthat bihday your material #model i love u take with u all the time urðy"±!!! ðy~™ðy~Žðy',,ðy'ðy'¦ðy factsguide: society now #motive tweet = combine['tidy_tweet'].agtweet.head()	selfish drags kids into dysfunction #run dysfunction #run thanks #lyft credit cause they offer wheelchair vans #disapointed #getthanked bihday your majesty #model love take with time ration factsguide society #motivation pply(lambda x: x.split())
	0 1 2 3 4	1 2 3 4 5 5 ooken	0.0 0.0 0.0 0.0 0.0	@user when a father is dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunction with the cause they don't offer wheelchair vans in #disapointed #getthat bihday your material with a little to be a selfished with a little with	when father dysfunctional selfish drags kids into dysfunction #run thanks #lyft credit cause they offer wheelchair vans #disapointed #getthanked bihday your majesty #model love take with time factsguide society #motivation
ut[15]:	0 1 2 3	1 2 3 4	0.0 0.0 0.0 0.0	@user when a father is dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his disaproinal into the disapoint and selfish he drags his disaproinal into his dysfunctional and selfish he drags his disaproinal into his dysfunctional and selfish he drags his dysfunctional and selfish he drags his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his dysfunctional and self-self-self-self-self-self-self-self-	is so when father dysfunctional selfish drags kids into dysfunction #run thanks #lyft credit cause they offer wheelchair vans #disapointed #getthanked bihday your majesty #model love take with time
ut[15]:	0 1 2	1 2 3	0.0	@user when a father is dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he disapointed in the disapoint of t	is so when father dysfunctional selfish drags kids into dysfunction #run dysfunction #run thanks #lyft credit cause they offer wheelchair vans anked #disapointed #getthanked bihday your majesty #me in #model love take with time
ut[15]:	0	1	0.0	@user when a father is dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish his disapointed and selfish his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he disappointed and selfish his disa	is so when father dysfunctional selfish drags kids into dysfunction #run thanks #lyft credit cause they offer wheelchair vans anked #disapointed #getthanked bijesty bihday your majesty
ut[15]:	0	1	0.0	@user when a father is dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his kids into his dysfunctional and selfish he drags his high he drags his high high	is so when father dysfunctional selfish drags kids into dysfunction #run tuse thanks #lyft credit cause they offer wheelchair vans
ut[15]:				@user when a father is dysfunctional and	is so when father dysfunctional ction. selfish drags kids into
ut[15]:		id	label	t	weet tidy_tweet
n [15]:			ne['t	<pre>#gr8! idy_tweet'] = combine['tidy_twee ad()</pre>	
	9	10	0.0	@user @user welcome here!i'm it's so	welcome here i m it s so #gr
	8	9	0.0	we won!!! love the land!!! #allin #cavs #champions #cleveland #clevelandcavaliers …	we won love the land #allin #cavs #champions #cleveland #clevelandcavaliers
	7	8	0.0	the next school year is the year for exams.ðŸ~- can't think about that ðŸ~- #school #exams #hate #imagine #actorslife #revolutionschool #girl	the next school year is the year for exams can t think about that #school #exams #hate #imagine #actorslife #revolutionschool #girl
	6	7	0.0	@user camping tomorrow @user @user @user @user @user @user danny…	camping tomorrow danny
	5	6	0.0	before they leave. chaos and pay disputes when they get there. #allshowandnogo	they leave chaos and pay disputes when they get there #allshowandnogo
	4	5	0.0	factsguide: society now #motivation [2/2] huge fan fare and big talking	factsguide society now #motivation huge fan fare and big talking before
	3	4	0.0	ur📱!!! ðŸ~™ðŸ~ŽðŸ′"ðŸ′…ðŸ′¦ðŸ′¦ðŸ′¦	time in ur
		3	0.0	bihday your majesty #model i love u take with u all the time in	bihday your majesty #model i love u take with u all the
	2			wheelchair vans in pdx. #disapointed #getthanked	vans in pdx #disapointed #getthanked
	1	2	0.0		

```
e, take, with, time]
                                                                      [factsguide, s
         ociety, #motivation]
         Name: tidy_tweet, dtype: object
In [17]:
          from nltk.stem.porter import *
          Stemmer = PorterStemmer()
          tokenized_tweet = tokenized_tweet.apply(lambda x: [Stemmer.stem(i) for i
In [18]:
          tokenized tweet
                                                               [when, father, dysfun
Out[18]:
         ct, selfish, drag, kid, into, dysfunct, #run]
                                               [thank, #lyft, credit, caus, they, o
         ffer, wheelchair, van, #disapoint, #getthank]
         [bihday, your, majesti]
         [#model, love, take, with, time]
         [factsguid, societi, #motiv]
                                  [thought, factori, left, right, polaris, #trump,
         49154
         #uselect, #leadership, #polit, #brexit, #blm]
                                           [feel, like, mermaid, #hairflip, #neverr
         eadi, #formal, #wed, #gown, #dress, #mermaid]
         49156
                  [#hillari, #campaign, today, #ohio, use, word, like, asset, liab
         il, never, onc, #clinton, thee, word, #radic]
                                            [happi, work, confer, right, mindset, l
         ead, cultur, develop, organ, #work, #mindset]
         49158
                                                                        [song, glad,
         free, download, #shoegaz, #newmus, #newsong]
         Name: tidy tweet, Length: 49159, dtype: object
In [19]:
          import nltk
          nltk.download('omw-1.4')
          from nltk.corpus import wordnet
          from nltk.stem.wordnet import WordNetLemmatizer
          Lemmatizer = WordNetLemmatizer()
          tokenized tweet = tokenized tweet.apply(lambda x: [Lemmatizer.lemmatize(j
         [nltk data] Downloading package omw-1.4 to
                         C:\Users\rsaipranayreddy\AppData\Roaming\nltk data...
         [nltk data]
         [nltk data]
                       Package omw-1.4 is already up-to-date!
In [20]:
          tokenized tweet
                                                              [when, father, dysfun
Out[20]:
         ct, selfish, drag, kid, into, dysfunct, #run]
                                               [thank, #lyft, credit, caus, they, o
         ffer, wheelchair, van, #disapoint, #getthank]
         [bihday, your, majesti]
```

```
[#model, love, take, with, time]
         [factsguid, societi, #motiv]
                                  [thought, factori, left, right, polaris, #trump,
         #uselect, #leadership, #polit, #brexit, #blm]
         49155
                                           [feel, like, mermaid, #hairflip, #neverr
         eadi, #formal, #wed, #gown, #dress, #mermaid]
                  [#hillari, #campaign, today, #ohio, use, word, like, asset, liab
         il, never, onc, #clinton, thee, word, #radic]
                                            [happi, work, confer, right, mindset, 1
         ead, cultur, develop, organ, #work, #mindset]
         49158
                                                                        [song, glad,
         free, download, #shoegaz, #newmus, #newsong]
         Name: tidy tweet, Length: 49159, dtype: object
In [21]:
          for i in range(len(tokenized tweet)):
              tokenized_tweet[i] = ' '.join(tokenized_tweet[i])
          combine['tidy_tweet'] = tokenized_tweet
In [22]:
          tokenized tweet
                                                        when father dysfunct selfis
Out[22]:
         h drag kid into dysfunct #run
                                          thank #lyft credit caus they offer wheelc
         hair van #disapoint #getthank
         bihday your majesti
         #model love take with time
         factsquid societi #motiv
                              thought factori left right polaris #trump #uselect #1
         eadership #polit #brexit #blm
                                      feel like mermaid #hairflip #neverreadi #form
         49155
         al #wed #gown #dress #mermaid
         49156
                  #hillari #campaign today #ohio use word like asset liabil never
         onc #clinton thee word #radic
         49157
                                        happi work confer right mindset lead cultur
         develop organ #work #mindset
         49158
                                                                song glad free downl
         oad #shoegaz #newmus #newsong
         Name: tidy tweet, Length: 49159, dtype: object
In [23]:
          combine
Out [23]:
                   id label
                                                       tweet
                                                                         tidy tweet
```

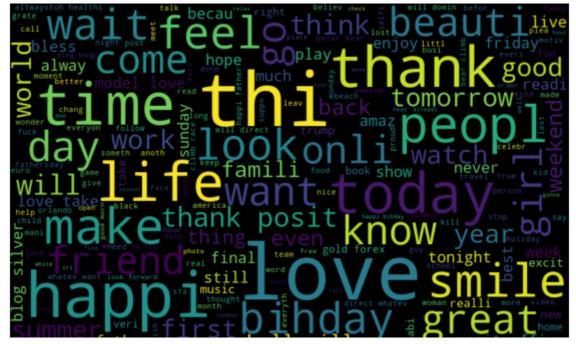
tidy_tweet	ei tweet		Iu		Uu L [23]
when father dysfunct selfish drag kid into dysfunct #run	@user when a father is dysfunctional and is so selfish he drags his kids into his dysfunction. #run	0.0	1	0	
thank #lyft credit caus they offer wheelchair van #disapoint #getthank	@user @user thanks for #lyft credit i can't use cause they don't offer wheelchair vans in pdx. #disapointed #getthanked	0.0	2	1	
bihdav vour maiesti	bihdav vour maiestv	0.0	3	2	

3	4	0.0	#model i love u take with u all the time in urðŸ"±!!! ðŸ~™ðŸ~ŽðŸ'"ðŸ'ðŸ'¦ðŸ'¦	#model love take with time
4	5	0.0	factsguide: society now #motivation	factsguid societi #motiv
•••				
49154	49155	NaN	thought factory: left-right polarisation! #trump #uselections2016 #leadership #politics #brexit #blm >3	thought factori left right polaris #trump #uselect #leadership #polit #brexit #blm
49155	49156	NaN	feeling like a mermaid ðŸ~~ #hairflip #neverready #formal #wedding #gown #dresses #mermaid …	feel like mermaid #hairflip #neverreadi #formal #wed #gown #dress #mermaid
49156	49157	NaN	#hillary #campaigned today in #ohio((omg)) & used words like "assets&liability" never once did #clinton say thee(word) #radicalization	#hillari #campaign today #ohio use word like asset liabil never onc #clinton thee word #radic
49157	49158	NaN	happy, at work conference: right mindset leads to culture-of-development organizations #work #mindset	happi work confer right mindset lead cultur develop organ #work #mindset
49158	49159	NaN	my song "so glad" free download! #shoegaze #newmusic #newsong	song glad free download #shoegaz #newmus #newsong

 $49159 \text{ rows} \times 4 \text{ columns}$

```
In [24]:
```

```
from wordcloud import WordCloud
all_words = ' '.join([text for text in combine['tidy_tweet']])
wordcloud = WordCloud(width= 800, height = 500, random_state = 21, max_fc
plt.figure(figsize=(10,7))
plt.imshow(wordcloud, interpolation = "bilinear")
plt.axis('off')
plt.show()
```



```
happen hate father silver gold bull WIII saturday share posit attirm
```

```
In [25]:
    normal_words = ' '.join([text for text in combine['tidy_tweet'][combine['
    wordcloud = WordCloud(width= 800, height = 500, random_state = 21, max_fc
    plt.figure(figsize=(10,7))
    plt.imshow(wordcloud, interpolation = "bilinear")
    plt.axis('off')
    plt.show()
```

```
Sunday

Tool kill

Too
```

```
In [26]:
    negative_words = ' '.join([text for text in combine['tidy_tweet'][combine
    wordcloud = WordCloud(width= 800, height = 500, random_state = 21, max_fc
    plt.figure(figsize=(10,7))
    plt.imshow(wordcloud, interpolation = "bilinear")
    plt.axis('off')
    plt.show()
```

```
Stop hatr calls WOMA Stand templ vandalis calgari wso truth plea better think templ vandalis calgari wso truth better think templ vandalis calgari wso truth plea better think templ vandalis calgari wso truth pl
```

Willies Case I all SO Latter SJW state liber of think black feel listen retweet time ignor think black feel listen retweet.

```
In [27]:
           def hashtag_extract(x):
               hashtags = []
               for i in x:
                    ht = re.findall(r"#(\w+)", i)
                    hashtags.append(ht)
               return hashtags
In [28]:
           HT_regular = hashtag_extract(combine['tidy_tweet'][combine['label'] == 0]
           HT_regular = sum(HT_regular,[])
In [29]:
           HT_negative = hashtag_extract(combine['tidy_tweet'][combine['label'] == 1
           HT_negative = sum(HT_negative,[])
In [30]:
           a = nltk.FreqDist(HT_regular)
           d = pd.DataFrame({'Hashtag': list(a.keys()), 'Count': list(a.values())})
           d = d.nlargest(columns = "Count", n = 20)
           plt.figure(figsize = (16,5))
           ax = sns.barplot(data = d, x = "Hashtag", y = "Count")
           ax.set(ylabel = 'Count')
           plt.show()
           1600
           1400
           1200
           1000
           800
            600
            400
            200
                  posit
                      smile healthi thank
                                  fun
                                         affirm summer model cute
                                                        beauti happi blog friend fathersday gold
                                                  Hashtag
In [31]:
           b = nltk.FreqDist(HT negative)
           d = pd.DataFrame({'Hashtag': list(b.keys()), 'Count': list(b.values())})
           d = d.nlargest(columns = "Count", n = 20)
           plt.figure(figsize = (16,5))
           ax = sns.barplot(data = d, x = "Hashtag", y = "Count")
           ax.set(ylabel = 'Count')
           plt.show()
           140
           120
           100
```