wrangle_act

December 13, 2019

0.1 Gathering and Extracting Data:

in this section we will collect data of WeRateDogs from Twitter using Tweepy API and tweets ids provided in twitter-archive-enhanced.csv file

In [1]: *#imporing all needed libraries*

```python
import tweepy
import requests as req
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import json
from io import BytesIO
import seaborn as sns
import re

%matplotlib inline
```

Tweepy API authorization

In [2]: *#providing all authorization keys to connect with Tweepy API*

```python
consumer_key = 'Hidden'
consumer_secret = 'Hidden'
access_token = 'Hidden'
access_secret = 'Hidden'

auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_secret)

api = tweepy.API(auth)
```

1 testing Api connection and extracting one tweet

tweet = api.get_status('1198860382976106496', tweet_mode='extended') print(tweet.retweet_count) print(tweet.favorite_count) print(tweet.full_text)

```
In [3]: #reading twitter-archive-enhanced.csv in pandas
        twitter_archive_df = pd.read_csv('twitter-archive-enhanced.csv')

In [4]: twitter_archive_df

Out[4]:                tweet_id  in_reply_to_status_id  in_reply_to_user_id  \
        0     892420643555336193                    NaN                  NaN
        1     892177421306343426                    NaN                  NaN
        2     891815181378084864                    NaN                  NaN
        3     891689557279858688                    NaN                  NaN
        4     891327558926688256                    NaN                  NaN
        5     891087950875897856                    NaN                  NaN
        6     890971913173991426                    NaN                  NaN
        7     890729181411237888                    NaN                  NaN
        8     890609185150312448                    NaN                  NaN
        9     890240255349198849                    NaN                  NaN
        10    890006608113172480                    NaN                  NaN
        11    889880896479866881                    NaN                  NaN
        12    889665388333682689                    NaN                  NaN
        13    889638837579907072                    NaN                  NaN
        14    889531135344209921                    NaN                  NaN
        15    889278841981685760                    NaN                  NaN
        16    888917238123831296                    NaN                  NaN
        17    888804989199671297                    NaN                  NaN
        18    888554962724278272                    NaN                  NaN
        19    888202515573088257                    NaN                  NaN
        20    888078434458587136                    NaN                  NaN
        21    887705289381826560                    NaN                  NaN
        22    887517139158093824                    NaN                  NaN
        23    887473957103951883                    NaN                  NaN
        24    887343217045368832                    NaN                  NaN
        25    887101392804085760                    NaN                  NaN
        26    886983233522544640                    NaN                  NaN
        27    886736880519319552                    NaN                  NaN
        28    886680336477933568                    NaN                  NaN
        29    886366144734445568                    NaN                  NaN
        ...                  ...                    ...                  ...
        2326  666411507551481857                    NaN                  NaN
        2327  666407126856765440                    NaN                  NaN
        2328  666396247373291520                    NaN                  NaN
        2329  666373753744588802                    NaN                  NaN
        2330  666362758909284353                    NaN                  NaN
        2331  666353288456101888                    NaN                  NaN
        2332  666345417576210432                    NaN                  NaN
        2333  666337882303524864                    NaN                  NaN
        2334  666293911632134144                    NaN                  NaN
        2335  666287406224695296                    NaN                  NaN
        2336  666273097616637952                    NaN                  NaN
```

2

```
2337  666268910803644416                          NaN          NaN
2338  666104133288665088                          NaN          NaN
2339  666102155909144576                          NaN          NaN
2340  666099513787052032                          NaN          NaN
2341  666094000022159362                          NaN          NaN
2342  666082916733198337                          NaN          NaN
2343  666073100786774016                          NaN          NaN
2344  666071193221509120                          NaN          NaN
2345  666063827256086533                          NaN          NaN
2346  666058600524156928                          NaN          NaN
2347  666057090499244032                          NaN          NaN
2348  666055525042405380                          NaN          NaN
2349  666051853826850816                          NaN          NaN
2350  666050758794694657                          NaN          NaN
2351  666049248165822465                          NaN          NaN
2352  666044226329800704                          NaN          NaN
2353  666033412701032449                          NaN          NaN
2354  666029285002620928                          NaN          NaN
2355  666020888022790149                          NaN          NaN

                     timestamp  \
0      2017-08-01 16:23:56 +0000
1      2017-08-01 00:17:27 +0000
2      2017-07-31 00:18:03 +0000
3      2017-07-30 15:58:51 +0000
4      2017-07-29 16:00:24 +0000
5      2017-07-29 00:08:17 +0000
6      2017-07-28 16:27:12 +0000
7      2017-07-28 00:22:40 +0000
8      2017-07-27 16:25:51 +0000
9      2017-07-26 15:59:51 +0000
10     2017-07-26 00:31:25 +0000
11     2017-07-25 16:11:53 +0000
12     2017-07-25 01:55:32 +0000
13     2017-07-25 00:10:02 +0000
14     2017-07-24 17:02:04 +0000
15     2017-07-24 00:19:32 +0000
16     2017-07-23 00:22:39 +0000
17     2017-07-22 16:56:37 +0000
18     2017-07-22 00:23:06 +0000
19     2017-07-21 01:02:36 +0000
20     2017-07-20 16:49:33 +0000
21     2017-07-19 16:06:48 +0000
22     2017-07-19 03:39:09 +0000
23     2017-07-19 00:47:34 +0000
24     2017-07-18 16:08:03 +0000
25     2017-07-18 00:07:08 +0000
26     2017-07-17 16:17:36 +0000
```

```
27      2017-07-16 23:58:41 +0000
28      2017-07-16 20:14:00 +0000
29      2017-07-15 23:25:31 +0000
...                          ...
2326    2015-11-17 00:24:19 +0000
2327    2015-11-17 00:06:54 +0000
2328    2015-11-16 23:23:41 +0000
2329    2015-11-16 21:54:18 +0000
2330    2015-11-16 21:10:36 +0000
2331    2015-11-16 20:32:58 +0000
2332    2015-11-16 20:01:42 +0000
2333    2015-11-16 19:31:45 +0000
2334    2015-11-16 16:37:02 +0000
2335    2015-11-16 16:11:11 +0000
2336    2015-11-16 15:14:19 +0000
2337    2015-11-16 14:57:41 +0000
2338    2015-11-16 04:02:55 +0000
2339    2015-11-16 03:55:04 +0000
2340    2015-11-16 03:44:34 +0000
2341    2015-11-16 03:22:39 +0000
2342    2015-11-16 02:38:37 +0000
2343    2015-11-16 01:59:36 +0000
2344    2015-11-16 01:52:02 +0000
2345    2015-11-16 01:22:45 +0000
2346    2015-11-16 01:01:59 +0000
2347    2015-11-16 00:55:59 +0000
2348    2015-11-16 00:49:46 +0000
2349    2015-11-16 00:35:11 +0000
2350    2015-11-16 00:30:50 +0000
2351    2015-11-16 00:24:50 +0000
2352    2015-11-16 00:04:52 +0000
2353    2015-11-15 23:21:54 +0000
2354    2015-11-15 23:05:30 +0000
2355    2015-11-15 22:32:08 +0000

                                                    source  \
0       <a href="http://twitter.com/download/iphone" r...
1       <a href="http://twitter.com/download/iphone" r...
2       <a href="http://twitter.com/download/iphone" r...
3       <a href="http://twitter.com/download/iphone" r...
4       <a href="http://twitter.com/download/iphone" r...
5       <a href="http://twitter.com/download/iphone" r...
6       <a href="http://twitter.com/download/iphone" r...
7       <a href="http://twitter.com/download/iphone" r...
8       <a href="http://twitter.com/download/iphone" r...
9       <a href="http://twitter.com/download/iphone" r...
10      <a href="http://twitter.com/download/iphone" r...
11      <a href="http://twitter.com/download/iphone" r...
```

```
12     <a href="http://twitter.com/download/iphone" r...
13     <a href="http://twitter.com/download/iphone" r...
14     <a href="http://twitter.com/download/iphone" r...
15     <a href="http://twitter.com/download/iphone" r...
16     <a href="http://twitter.com/download/iphone" r...
17     <a href="http://twitter.com/download/iphone" r...
18     <a href="http://twitter.com/download/iphone" r...
19     <a href="http://twitter.com/download/iphone" r...
20     <a href="http://twitter.com/download/iphone" r...
21     <a href="http://twitter.com/download/iphone" r...
22     <a href="http://twitter.com/download/iphone" r...
23     <a href="http://twitter.com/download/iphone" r...
24     <a href="http://twitter.com/download/iphone" r...
25     <a href="http://twitter.com/download/iphone" r...
26     <a href="http://twitter.com/download/iphone" r...
27     <a href="http://twitter.com/download/iphone" r...
28     <a href="http://twitter.com/download/iphone" r...
29     <a href="http://twitter.com/download/iphone" r...
...                                                  ...
2326   <a href="http://twitter.com/download/iphone" r...
2327   <a href="http://twitter.com/download/iphone" r...
2328   <a href="http://twitter.com/download/iphone" r...
2329   <a href="http://twitter.com/download/iphone" r...
2330   <a href="http://twitter.com/download/iphone" r...
2331   <a href="http://twitter.com/download/iphone" r...
2332   <a href="http://twitter.com/download/iphone" r...
2333   <a href="http://twitter.com/download/iphone" r...
2334   <a href="http://twitter.com/download/iphone" r...
2335   <a href="http://twitter.com/download/iphone" r...
2336   <a href="http://twitter.com/download/iphone" r...
2337   <a href="http://twitter.com/download/iphone" r...
2338   <a href="http://twitter.com/download/iphone" r...
2339   <a href="http://twitter.com/download/iphone" r...
2340   <a href="http://twitter.com/download/iphone" r...
2341   <a href="http://twitter.com/download/iphone" r...
2342   <a href="http://twitter.com/download/iphone" r...
2343   <a href="http://twitter.com/download/iphone" r...
2344   <a href="http://twitter.com/download/iphone" r...
2345   <a href="http://twitter.com/download/iphone" r...
2346   <a href="http://twitter.com/download/iphone" r...
2347   <a href="http://twitter.com/download/iphone" r...
2348   <a href="http://twitter.com/download/iphone" r...
2349   <a href="http://twitter.com/download/iphone" r...
2350   <a href="http://twitter.com/download/iphone" r...
2351   <a href="http://twitter.com/download/iphone" r...
2352   <a href="http://twitter.com/download/iphone" r...
2353   <a href="http://twitter.com/download/iphone" r...
2354   <a href="http://twitter.com/download/iphone" r...
```

```
2355  <a href="http://twitter.com/download/iphone" r...

                                               text  retweeted_status_id  \
0     This is Phineas. He's a mystical boy. Only eve...                  NaN
1     This is Tilly. She's just checking pup on you...                   NaN
2     This is Archie. He is a rare Norwegian Pouncin...                  NaN
3     This is Darla. She commenced a snooze mid meal...                  NaN
4     This is Franklin. He would like you to stop ca...                  NaN
5     Here we have a majestic great white breaching ...                  NaN
6     Meet Jax. He enjoys ice cream so much he gets ...                  NaN
7     When you watch your owner call another dog a g...                  NaN
8     This is Zoey. She doesn't want to be one of th...                  NaN
9     This is Cassie. She is a college pup. Studying...                  NaN
10    This is Koda. He is a South Australian decksha...                  NaN
11    This is Bruno. He is a service shark. Only get...                  NaN
12    Here's a puppo that seems to be on the fence a...                  NaN
13    This is Ted. He does his best. Sometimes that'...                  NaN
14    This is Stuart. He's sporting his favorite fan...                  NaN
15    This is Oliver. You're witnessing one of his m...                  NaN
16    This is Jim. He found a fren. Taught him how t...                  NaN
17    This is Zeke. He has a new stick. Very proud o...                  NaN
18    This is Ralphus. He's powering up. Attempting ...                  NaN
19    RT @dog_rates: This is Canela. She attempted s...         8.874740e+17
20    This is Gerald. He was just told he didn't get...                  NaN
21    This is Jeffrey. He has a monopoly on the pool...                  NaN
22    I've yet to rate a Venezuelan Hover Wiener. Th...                  NaN
23    This is Canela. She attempted some fancy porch...                  NaN
24    You may not have known you needed to see this ...                  NaN
25    This... is a Jubilant Antarctic House Bear. We...                  NaN
26    This is Maya. She's very shy. Rarely leaves he...                  NaN
27    This is Mingus. He's a wonderful father to his...                  NaN
28    This is Derek. He's late for a dog meeting. 13...                  NaN
29    This is Roscoe. Another pupper fallen victim t...                  NaN
...                                                 ...                  ...
2326  This is quite the dog. Gets really excited whe...                  NaN
2327  This is a southern Vesuvius bumblegruff. Can d...                  NaN
2328  Oh goodness. A super rare northeast Qdoba kang...                  NaN
2329  Those are sunglasses and a jean jacket. 11/10 ...                  NaN
2330  Unique dog here. Very small. Lives in containe...                  NaN
2331  Here we have a mixed Asiago from the Galápagos...                  NaN
2332  Look at this jokester thinking seat belt laws ...                  NaN
2333  This is an extremely rare horned Parthenon. No...                  NaN
2334  This is a funny dog. Weird toes. Won't come do...                  NaN
2335  This is an Albanian 3 1/2 legged  Episcopalian...                  NaN
2336       Can take selfies 11/10 https://t.co/ws2AMaNwPW                 NaN
2337  Very concerned about fellow dog trapped in com...                  NaN
2338  Not familiar with this breed. No tail (weird)...                   NaN
2339  Oh my. Here you are seeing an Adobe Setter giv...                  NaN
```

```
2340  Can stand on stump for what seems like a while...          NaN
2341  This appears to be a Mongolian Presbyterian mi...          NaN
2342  Here we have a well-established sunblockerspan...          NaN
2343  Let's hope this flight isn't Malaysian (lol). ...          NaN
2344  Here we have a northern speckled Rhododendron...     NaN
2345  This is the happiest dog you will ever see. Ve...          NaN
2346  Here is the Rand Paul of retrievers folks! He'...          NaN
2347  My oh my. This is a rare blond Canadian terrie...          NaN
2348  Here is a Siberian heavily armored polar bear ...          NaN
2349  This is an odd dog. Hard on the outside but lo...          NaN
2350  This is a truly beautiful English Wilson Staff...          NaN
2351  Here we have a 1949 1st generation vulpix. Enj...          NaN
2352  This is a purebred Piers Morgan. Loves to Netf...          NaN
2353  Here is a very happy pup. Big fan of well-main...          NaN
2354  This is a western brown Mitsubishi terrier. Up...          NaN
2355  Here we have a Japanese Irish Setter. Lost eye...          NaN

      retweeted_status_user_id retweeted_status_timestamp  \
0                          NaN                        NaN
1                          NaN                        NaN
2                          NaN                        NaN
3                          NaN                        NaN
4                          NaN                        NaN
5                          NaN                        NaN
6                          NaN                        NaN
7                          NaN                        NaN
8                          NaN                        NaN
9                          NaN                        NaN
10                         NaN                        NaN
11                         NaN                        NaN
12                         NaN                        NaN
13                         NaN                        NaN
14                         NaN                        NaN
15                         NaN                        NaN
16                         NaN                        NaN
17                         NaN                        NaN
18                         NaN                        NaN
19                4.196984e+09  2017-07-19 00:47:34 +0000
20                         NaN                        NaN
21                         NaN                        NaN
22                         NaN                        NaN
23                         NaN                        NaN
24                         NaN                        NaN
25                         NaN                        NaN
26                         NaN                        NaN
27                         NaN                        NaN
28                         NaN                        NaN
29                         NaN                        NaN
```

```
...                      ...                    ...
2326                     NaN                    NaN
2327                     NaN                    NaN
2328                     NaN                    NaN
2329                     NaN                    NaN
2330                     NaN                    NaN
2331                     NaN                    NaN
2332                     NaN                    NaN
2333                     NaN                    NaN
2334                     NaN                    NaN
2335                     NaN                    NaN
2336                     NaN                    NaN
2337                     NaN                    NaN
2338                     NaN                    NaN
2339                     NaN                    NaN
2340                     NaN                    NaN
2341                     NaN                    NaN
2342                     NaN                    NaN
2343                     NaN                    NaN
2344                     NaN                    NaN
2345                     NaN                    NaN
2346                     NaN                    NaN
2347                     NaN                    NaN
2348                     NaN                    NaN
2349                     NaN                    NaN
2350                     NaN                    NaN
2351                     NaN                    NaN
2352                     NaN                    NaN
2353                     NaN                    NaN
2354                     NaN                    NaN
2355                     NaN                    NaN

                                        expanded_urls  rating_numerator  \
0      https://twitter.com/dog_rates/status/892420643...                13
1      https://twitter.com/dog_rates/status/892177421...                13
2      https://twitter.com/dog_rates/status/891815181...                12
3      https://twitter.com/dog_rates/status/891689557...                13
4      https://twitter.com/dog_rates/status/891327558...                12
5      https://twitter.com/dog_rates/status/891087950...                13
6      https://gofundme.com/ydvmve-surgery-for-jax,ht...                13
7      https://twitter.com/dog_rates/status/890729181...                13
8      https://twitter.com/dog_rates/status/890609185...                13
9      https://twitter.com/dog_rates/status/890240255...                14
10     https://twitter.com/dog_rates/status/890006608...                13
11     https://twitter.com/dog_rates/status/889880896...                13
12     https://twitter.com/dog_rates/status/889665388...                13
13     https://twitter.com/dog_rates/status/889638837...                12
14     https://twitter.com/dog_rates/status/889531135...                13
```

8

```
15     https://twitter.com/dog_rates/status/889278841...              13
16     https://twitter.com/dog_rates/status/888917238...              12
17     https://twitter.com/dog_rates/status/888804989...              13
18     https://twitter.com/dog_rates/status/888554962...              13
19     https://twitter.com/dog_rates/status/887473957...              13
20     https://twitter.com/dog_rates/status/888078434...              12
21     https://twitter.com/dog_rates/status/887705289...              13
22     https://twitter.com/dog_rates/status/887517139...              14
23     https://twitter.com/dog_rates/status/887473957...              13
24     https://twitter.com/dog_rates/status/887343217...              13
25     https://twitter.com/dog_rates/status/887101392...              12
26     https://twitter.com/dog_rates/status/886983233...              13
27     https://www.gofundme.com/mingusneedsus,https:/...              13
28     https://twitter.com/dog_rates/status/886680336...              13
29     https://twitter.com/dog_rates/status/886366144...              12
...                                                 ...              ...
2326   https://twitter.com/dog_rates/status/666411507...               2
2327   https://twitter.com/dog_rates/status/666407126...               7
2328   https://twitter.com/dog_rates/status/666396247...               9
2329   https://twitter.com/dog_rates/status/666373753...              11
2330   https://twitter.com/dog_rates/status/666362758...               6
2331   https://twitter.com/dog_rates/status/666353288...               8
2332   https://twitter.com/dog_rates/status/666345417...              10
2333   https://twitter.com/dog_rates/status/666337882...               9
2334   https://twitter.com/dog_rates/status/666293911...               3
2335   https://twitter.com/dog_rates/status/666287406...               1
2336   https://twitter.com/dog_rates/status/666273097...              11
2337   https://twitter.com/dog_rates/status/666268910...              10
2338   https://twitter.com/dog_rates/status/666104133...               1
2339   https://twitter.com/dog_rates/status/666102155...              11
2340   https://twitter.com/dog_rates/status/666099513...               8
2341   https://twitter.com/dog_rates/status/666094000...               9
2342   https://twitter.com/dog_rates/status/666082916...               6
2343   https://twitter.com/dog_rates/status/666073100...              10
2344   https://twitter.com/dog_rates/status/666071193...               9
2345   https://twitter.com/dog_rates/status/666063827...              10
2346   https://twitter.com/dog_rates/status/666058600...               8
2347   https://twitter.com/dog_rates/status/666057090...               9
2348   https://twitter.com/dog_rates/status/666055525...              10
2349   https://twitter.com/dog_rates/status/666051853...               2
2350   https://twitter.com/dog_rates/status/666050758...              10
2351   https://twitter.com/dog_rates/status/666049248...               5
2352   https://twitter.com/dog_rates/status/666044226...               6
2353   https://twitter.com/dog_rates/status/666033412...               9
2354   https://twitter.com/dog_rates/status/666029285...               7
2355   https://twitter.com/dog_rates/status/666020888...               8

       rating_denominator      name  doggo floofer  pupper  puppo
```

0	10	Phineas	None	None	None	None
1	10	Tilly	None	None	None	None
2	10	Archie	None	None	None	None
3	10	Darla	None	None	None	None
4	10	Franklin	None	None	None	None
5	10	None	None	None	None	None
6	10	Jax	None	None	None	None
7	10	None	None	None	None	None
8	10	Zoey	None	None	None	None
9	10	Cassie	doggo	None	None	None
10	10	Koda	None	None	None	None
11	10	Bruno	None	None	None	None
12	10	None	None	None	None	puppo
13	10	Ted	None	None	None	None
14	10	Stuart	None	None	None	puppo
15	10	Oliver	None	None	None	None
16	10	Jim	None	None	None	None
17	10	Zeke	None	None	None	None
18	10	Ralphus	None	None	None	None
19	10	Canela	None	None	None	None
20	10	Gerald	None	None	None	None
21	10	Jeffrey	None	None	None	None
22	10	such	None	None	None	None
23	10	Canela	None	None	None	None
24	10	None	None	None	None	None
25	10	None	None	None	None	None
26	10	Maya	None	None	None	None
27	10	Mingus	None	None	None	None
28	10	Derek	None	None	None	None
29	10	Roscoe	None	None	pupper	None
...	...	...	...	...	...	...
2326	10	quite	None	None	None	None
2327	10	a	None	None	None	None
2328	10	None	None	None	None	None
2329	10	None	None	None	None	None
2330	10	None	None	None	None	None
2331	10	None	None	None	None	None
2332	10	None	None	None	None	None
2333	10	an	None	None	None	None
2334	10	a	None	None	None	None
2335	2	an	None	None	None	None
2336	10	None	None	None	None	None
2337	10	None	None	None	None	None
2338	10	None	None	None	None	None
2339	10	None	None	None	None	None
2340	10	None	None	None	None	None
2341	10	None	None	None	None	None
2342	10	None	None	None	None	None