

wrangle_act

October 21, 2020

```
In [8]: import pandas as pd
import numpy as np
import requests
import os
import tweepy
import json
import re
from IPython.display import HTML
```

```
In [9]: weratedogs_df = pd.read_csv('twitter-archive-enhanced.csv')
```

```
In [10]: url = 'https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad_image-predicti
tsv_file = requests.get(url)
tsv_file
```

```
Out[10]: <Response [200]>
```

```
In [11]: weratedogs_df
```

```
Out[11]:
```

| | 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 |
| 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/ws2AMaWpPW | 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 | |
| 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 |
| 2343 | 10 | None | None | None | None | None |
| 2344 | 10 | None | None | None | None | None |
| 2345 | 10 | the | None | None | None | None |
| 2346 | 10 | the | None | None | None | None |
| 2347 | 10 | a | None | None | None | None |
| 2348 | 10 | a | None | None | None | None |
| 2349 | 10 | an | None | None | None | None |
| 2350 | 10 | a | None | None | None | None |
| 2351 | 10 | None | None | None | None | None |
| 2352 | 10 | a | None | None | None | None |
| 2353 | 10 | a | None | None | None | None |
| 2354 | 10 | a | None | None | None | None |
| 2355 | 10 | None | None | None | None | None |

[2356 rows x 17 columns]

```
In [12]: tsv_folder = 'photos'
         if not os.path.exists(tsv_folder):
             os.makedirs(tsv_folder)

In [13]: file_name = url.split('/')[ -1]
         with open(file_name,mode = 'wb') as file:
             file.write(tsv_file.content)

In [14]: predictions_df = pd.read_csv('image-predictions.tsv', sep = '\t')
```

```

In [15]: import tweepy
         from tweepy import OAuthHandler
         import json
         from timeit import default_timer as timer

         # Query Twitter API for each tweet in the Twitter archive and save JSON in a text file
         # These are hidden to comply with Twitter's API terms and conditions
         consumer_key = 'HIDDEN'
         consumer_secret = 'HIDDEN'
         access_token = 'HIDDEN'
         access_secret = 'HIDDEN'

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

         api = tweepy.API(auth, wait_on_rate_limit=True)

         # NOTE TO STUDENT WITH MOBILE VERIFICATION ISSUES:
         # df_1 is a DataFrame with the twitter_archive_enhanced.csv file. You may have to
         # change line 17 to match the name of your DataFrame with twitter_archive_enhanced.csv
         # NOTE TO REVIEWER: this student had mobile verification issues so the following
         # Twitter API code was sent to this student from a Udacity instructor
         # Tweet IDs for which to gather additional data via Twitter's API
         tweet_ids = df_1.tweet_id.values
         len(tweet_ids)

         # Query Twitter's API for JSON data for each tweet ID in the Twitter archive
         count = 0
         fails_dict = {}
         start = timer()
         # Save each tweet's returned JSON as a new line in a .txt file
         with open('tweet_json.txt', 'w') as outfile:
             # This loop will likely take 20-30 minutes to run because of Twitter's rate limit
             for tweet_id in tweet_ids:
                 count += 1
                 print(str(count) + ": " + str(tweet_id))
                 try:
                     tweet = api.get_status(tweet_id, tweet_mode='extended')

```

```

        print("Success")
        json.dump(tweet._json, outfile)
        outfile.write('\n')
    except tweepy.TweepError as e:
        print("Fail")
        fails_dict[tweet_id] = e
        pass
end = timer()
print(end - start)
print(fails_dict)

```

NameError Traceback (most recent call last)

```

<ipython-input-15-144479a83cde> in <module>()
    30 # Twitter API code was sent to this student from a Udacity instructor
    31 # Tweet IDs for which to gather additional data via Twitter's API
---> 32 tweet_ids = df_1.tweet_id.values
    33 len(tweet_ids)
    34

```

NameError: name 'df_1' is not defined

In [16]: df_list = []

```

with open('tweet-json.txt', 'r') as file:
    for line in file:
        tweet = json.loads(line)
        tweet_id = tweet['id']
        retweet_count = tweet['retweet_count']
        fav_count = tweet['favorite_count']
        df_list.append({'tweet_id':tweet_id,
                        'retweet_count': retweet_count,
                        'favorite_count': fav_count})

api_df_now = pd.DataFrame(df_list)
api_df_now.head()

```

```

Out[16]:
   favorite_count  retweet_count  tweet_id
0           39467           8853  892420643555336193
1           33819           6514  892177421306343426
2           25461           4328  891815181378084864

```

| | | | |
|---|-------|------|--------------------|
| 3 | 42908 | 8964 | 891689557279858688 |
| 4 | 41048 | 9774 | 891327558926688256 |

1 Quality 1)"weratedogs_df" has 2356 rows while the "photos_df" has only 2075 rows, this maybe due to retweets or missing photos. 2)denominator is higher than 10 or equals 0 in some cases 3)numerator higher than 10 4)dog names are misspelled or written wrongly or not written at all 5)data types need amendments 6)weratedogs_df contains 181 retweets and 78 replies which are not applicable 7)the time stamp contains unnecessary things (+000) 8)reorder.

2 Tidiness 1)4 columns were found for doggo, floffer, pupper, and puppo in the "werate-dogs_df" but this is same categorical data that would be better in only 1 column. 2)3 predictions were observed for the dog breed in the "predictions_df" 3)3 data frames were used photos_df, weratedogs_df, and api_df_now

In [18]: predictions_df

```
Out[18]:
```

| | tweet_id | jpg_url \ |
|------|--------------------|---|
| 0 | 666020888022790149 | https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg |
| 1 | 666029285002620928 | https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg |
| 2 | 666033412701032449 | https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg |
| 3 | 666044226329800704 | https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg |
| 4 | 666049248165822465 | https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg |
| 5 | 666050758794694657 | https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg |
| 6 | 666051853826850816 | https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg |
| 7 | 666055525042405380 | https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg |
| 8 | 666057090499244032 | https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg |
| 9 | 666058600524156928 | https://pbs.twimg.com/media/CT5Qw94XAAA-2dP.jpg |
| 10 | 666063827256086533 | https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg |
| 11 | 666071193221509120 | https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg |
| 12 | 666073100786774016 | https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg |
| 13 | 666082916733198337 | https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg |
| 14 | 666094000022159362 | https://pbs.twimg.com/media/CT5w9gUW4AAsBNN.jpg |
| 15 | 666099513787052032 | https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg |
| 16 | 666102155909144576 | https://pbs.twimg.com/media/CT54YGiWUAENZnoK.jpg |
| 17 | 666104133288665088 | https://pbs.twimg.com/media/CT56LSZW0AALJj2.jpg |
| 18 | 666268910803644416 | https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg |
| 19 | 666273097616637952 | https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg |
| 20 | 666287406224695296 | https://pbs.twimg.com/media/CT8g3BpUEAAuFjg.jpg |
| 21 | 666293911632134144 | https://pbs.twimg.com/media/CT8mx7KW4AEQu8N.jpg |
| 22 | 666337882303524864 | https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg |
| 23 | 666345417576210432 | https://pbs.twimg.com/media/CT9Vn7PW0AA_ZCM.jpg |
| 24 | 666353288456101888 | https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg |
| 25 | 666362758909284353 | https://pbs.twimg.com/media/CT9lXGsUcAAyUft.jpg |
| 26 | 666373753744588802 | https://pbs.twimg.com/media/CT9vZEYWUAA1Z05.jpg |
| 27 | 666396247373291520 | https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg |
| 28 | 666407126856765440 | https://pbs.twimg.com/media/CT-NvwmW4AAugGZ.jpg |
| 29 | 666411507551481857 | https://pbs.twimg.com/media/CT-RugiWIAELEaq.jpg |
| ... | ... | ... |
| 2045 | 886366144734445568 | https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg |

| | | |
|------|--------------------|---|
| 2046 | 886680336477933568 | https://pbs.twimg.com/media/DE4fEDzWAAAYHMM.jpg |
| 2047 | 886736880519319552 | https://pbs.twimg.com/media/DE5Se8FXcAAJFx4.jpg |
| 2048 | 886983233522544640 | https://pbs.twimg.com/media/DE8yicJW0AAAvBJ.jpg |
| 2049 | 887101392804085760 | https://pbs.twimg.com/media/DE-eAq6UwAA-jaE.jpg |
| 2050 | 887343217045368832 | https://pbs.twimg.com/ext_tw_video_thumb/88734... |
| 2051 | 887473957103951883 | https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg |
| 2052 | 887517139158093824 | https://pbs.twimg.com/ext_tw_video_thumb/88751... |
| 2053 | 887705289381826560 | https://pbs.twimg.com/media/DFHDQBbXgAEqY7t.jpg |
| 2054 | 888078434458587136 | https://pbs.twimg.com/media/DFMWn56WsAAKA7B.jpg |
| 2055 | 888202515573088257 | https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg |
| 2056 | 888554962724278272 | https://pbs.twimg.com/media/DFTH_0-UQAACu20.jpg |
| 2057 | 888804989199671297 | https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg |
| 2058 | 888917238123831296 | https://pbs.twimg.com/media/DFYRgsOUQAARGh0.jpg |
| 2059 | 889278841981685760 | https://pbs.twimg.com/ext_tw_video_thumb/88927... |
| 2060 | 889531135344209921 | https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg |
| 2061 | 889638837579907072 | https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg |
| 2062 | 889665388333682689 | https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg |
| 2063 | 889880896479866881 | https://pbs.twimg.com/media/DF199B1WsAITKsg.jpg |
| 2064 | 890006608113172480 | https://pbs.twimg.com/media/DFnwSY4WAAAMliS.jpg |
| 2065 | 890240255349198849 | https://pbs.twimg.com/media/DFrEyVuW0AA03t9.jpg |
| 2066 | 890609185150312448 | https://pbs.twimg.com/media/DFwUU__XcAEpyXI.jpg |
| 2067 | 890729181411237888 | https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg |
| 2068 | 890971913173991426 | https://pbs.twimg.com/media/DF1eOmZXUAAALUcq.jpg |
| 2069 | 891087950875897856 | https://pbs.twimg.com/media/DF3HwyEWsAABqE6.jpg |
| 2070 | 891327558926688256 | https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg |
| 2071 | 891689557279858688 | https://pbs.twimg.com/media/DF_q7IAWsAEuuN8.jpg |
| 2072 | 891815181378084864 | https://pbs.twimg.com/media/DGBdLU1WsAANxJ9.jpg |
| 2073 | 892177421306343426 | https://pbs.twimg.com/media/DGGmoV4XsAAUL6n.jpg |
| 2074 | 892420643555336193 | https://pbs.twimg.com/media/DGKD1-bXoAAIAUK.jpg |

| | img_num | | p1 | p1_conf | p1_dog | \ |
|----|---------|------------------------|----------|---------|--------|---|
| 0 | 1 | Welsh_springer_spaniel | 0.465074 | | True | |
| 1 | 1 | redbone | 0.506826 | | True | |
| 2 | 1 | German_shepherd | 0.596461 | | True | |
| 3 | 1 | Rhodesian_ridgeback | 0.408143 | | True | |
| 4 | 1 | miniature_pinscher | 0.560311 | | True | |
| 5 | 1 | Bernese_mountain_dog | 0.651137 | | True | |
| 6 | 1 | box_turtle | 0.933012 | | False | |
| 7 | 1 | chow | 0.692517 | | True | |
| 8 | 1 | shopping_cart | 0.962465 | | False | |
| 9 | 1 | miniature_poodle | 0.201493 | | True | |
| 10 | 1 | golden_retriever | 0.775930 | | True | |
| 11 | 1 | Gordon_setter | 0.503672 | | True | |
| 12 | 1 | Walker_hound | 0.260857 | | True | |
| 13 | 1 | pug | 0.489814 | | True | |
| 14 | 1 | bloodhound | 0.195217 | | True | |
| 15 | 1 | Lhasa | 0.582330 | | True | |
| 16 | 1 | English_setter | 0.298617 | | True | |

| | | | | |
|------|-----|-----------------------------|----------|-------|
| 17 | 1 | hen | 0.965932 | False |
| 18 | 1 | desktop_computer | 0.086502 | False |
| 19 | 1 | Italian_greyhound | 0.176053 | True |
| 20 | 1 | Maltese_dog | 0.857531 | True |
| 21 | 1 | three-toed_sloth | 0.914671 | False |
| 22 | 1 | ox | 0.416669 | False |
| 23 | 1 | golden_retriever | 0.858744 | True |
| 24 | 1 | malamute | 0.336874 | True |
| 25 | 1 | guinea_pig | 0.996496 | False |
| 26 | 1 | soft-coated_wheaten_terrier | 0.326467 | True |
| 27 | 1 | Chihuahua | 0.978108 | True |
| 28 | 1 | black-and-tan_coonhound | 0.529139 | True |
| 29 | 1 | coho | 0.404640 | False |
| ... | ... | ... | ... | ... |
| 2045 | 1 | French_bulldog | 0.999201 | True |
| 2046 | 1 | convertible | 0.738995 | False |
| 2047 | 1 | kuvasz | 0.309706 | True |
| 2048 | 2 | Chihuahua | 0.793469 | True |
| 2049 | 1 | Samoyed | 0.733942 | True |
| 2050 | 1 | Mexican_hairless | 0.330741 | True |
| 2051 | 2 | Pembroke | 0.809197 | True |
| 2052 | 1 | limousine | 0.130432 | False |
| 2053 | 1 | basset | 0.821664 | True |
| 2054 | 1 | French_bulldog | 0.995026 | True |
| 2055 | 2 | Pembroke | 0.809197 | True |
| 2056 | 3 | Siberian_husky | 0.700377 | True |
| 2057 | 1 | golden_retriever | 0.469760 | True |
| 2058 | 1 | golden_retriever | 0.714719 | True |
| 2059 | 1 | whippet | 0.626152 | True |
| 2060 | 1 | golden_retriever | 0.953442 | True |
| 2061 | 1 | French_bulldog | 0.991650 | True |
| 2062 | 1 | Pembroke | 0.966327 | True |
| 2063 | 1 | French_bulldog | 0.377417 | True |
| 2064 | 1 | Samoyed | 0.957979 | True |
| 2065 | 1 | Pembroke | 0.511319 | True |
| 2066 | 1 | Irish_terrier | 0.487574 | True |
| 2067 | 2 | Pomeranian | 0.566142 | True |
| 2068 | 1 | Appenzeller | 0.341703 | True |
| 2069 | 1 | Chesapeake_Bay_retriever | 0.425595 | True |
| 2070 | 2 | basset | 0.555712 | True |
| 2071 | 1 | paper_towel | 0.170278 | False |
| 2072 | 1 | Chihuahua | 0.716012 | True |
| 2073 | 1 | Chihuahua | 0.323581 | True |
| 2074 | 1 | orange | 0.097049 | False |

| | | p2 | p2_conf | p2_dog | p3 \ |
|---|--------------------|----------|---------|---------------------|------|
| 0 | collie | 0.156665 | True | Shetland_sheepdog | |
| 1 | miniature_pinscher | 0.074192 | True | Rhodesian_ridgeback | |

| | | | | |
|------|--------------------------|----------|-------|-----------------------------|
| 2 | malinois | 0.138584 | True | bloodhound |
| 3 | redbone | 0.360687 | True | miniature_pinscher |
| 4 | Rottweiler | 0.243682 | True | Doberman |
| 5 | English_springer | 0.263788 | True | Greater_Swiss_Mountain_dog |
| 6 | mud_turtle | 0.045885 | False | terrapiin |
| 7 | Tibetan_mastiff | 0.058279 | True | fur_coat |
| 8 | shopping_basket | 0.014594 | False | golden_retriever |
| 9 | komondor | 0.192305 | True | soft-coated_wheaten_terrier |
| 10 | Tibetan_mastiff | 0.093718 | True | Labrador_retriever |
| 11 | Yorkshire_terrier | 0.174201 | True | Pekinese |
| 12 | English_foxhound | 0.175382 | True | Ibizan_hound |
| 13 | bull_mastiff | 0.404722 | True | French_bulldog |
| 14 | German_shepherd | 0.078260 | True | malinois |
| 15 | Shih-Tzu | 0.166192 | True | Dandie_Dinmont |
| 16 | Newfoundland | 0.149842 | True | borzoi |
| 17 | cock | 0.033919 | False | partridge |
| 18 | desk | 0.085547 | False | bookcase |
| 19 | toy_terrier | 0.111884 | True | basenji |
| 20 | toy_poodle | 0.063064 | True | miniature_poodle |
| 21 | otter | 0.015250 | False | great_grey_owl |
| 22 | Newfoundland | 0.278407 | True | groenendael |
| 23 | Chesapeake_Bay_retriever | 0.054787 | True | Labrador_retriever |
| 24 | Siberian_husky | 0.147655 | True | Eskimo_dog |
| 25 | skunk | 0.002402 | False | hamster |
| 26 | Afghan_hound | 0.259551 | True | briard |
| 27 | toy_terrier | 0.009397 | True | papillon |
| 28 | bloodhound | 0.244220 | True | flat-coated_retriever |
| 29 | barracouta | 0.271485 | False | gar |
| ... | ... | ... | ... | ... |
| 2045 | Chihuahua | 0.000361 | True | Boston_bull |
| 2046 | sports_car | 0.139952 | False | car_wheel |
| 2047 | Great_Pyrenees | 0.186136 | True | Dandie_Dinmont |
| 2048 | toy_terrier | 0.143528 | True | can_opener |
| 2049 | Eskimo_dog | 0.035029 | True | Staffordshire_bullterrier |
| 2050 | sea_lion | 0.275645 | False | Weimaraner |
| 2051 | Rhodesian_ridgeback | 0.054950 | True | beagle |
| 2052 | tow_truck | 0.029175 | False | shopping_cart |
| 2053 | redbone | 0.087582 | True | Weimaraner |
| 2054 | pug | 0.000932 | True | bull_mastiff |
| 2055 | Rhodesian_ridgeback | 0.054950 | True | beagle |
| 2056 | Eskimo_dog | 0.166511 | True | malamute |
| 2057 | Labrador_retriever | 0.184172 | True | English_setter |
| 2058 | Tibetan_mastiff | 0.120184 | True | Labrador_retriever |
| 2059 | borzoi | 0.194742 | True | Saluki |
| 2060 | Labrador_retriever | 0.013834 | True | redbone |
| 2061 | boxer | 0.002129 | True | Staffordshire_bullterrier |
| 2062 | Cardigan | 0.027356 | True | basenji |
| 2063 | Labrador_retriever | 0.151317 | True | muzzle |

| | | | | |
|------|--------------------|----------|-------|-----------------------------|
| 2064 | Pomeranian | 0.013884 | True | chow |
| 2065 | Cardigan | 0.451038 | True | Chihuahua |
| 2066 | Irish_setter | 0.193054 | True | Chesapeake_Bay_retriever |
| 2067 | Eskimo_dog | 0.178406 | True | Pembroke |
| 2068 | Border_collie | 0.199287 | True | ice_lolly |
| 2069 | Irish_terrier | 0.116317 | True | Indian_elephant |
| 2070 | English_springer | 0.225770 | True | German_short-haired_pointer |
| 2071 | Labrador_retriever | 0.168086 | True | spatula |
| 2072 | malamute | 0.078253 | True | kelpie |
| 2073 | Pekinese | 0.090647 | True | papillon |
| 2074 | bagel | 0.085851 | False | banana |

| | p3_conf | p3_dog |
|------|----------|--------|
| 0 | 0.061428 | True |
| 1 | 0.072010 | True |
| 2 | 0.116197 | True |
| 3 | 0.222752 | True |
| 4 | 0.154629 | True |
| 5 | 0.016199 | True |
| 6 | 0.017885 | False |
| 7 | 0.054449 | False |
| 8 | 0.007959 | True |
| 9 | 0.082086 | True |
| 10 | 0.072427 | True |
| 11 | 0.109454 | True |
| 12 | 0.097471 | True |
| 13 | 0.048960 | True |
| 14 | 0.075628 | True |
| 15 | 0.089688 | True |
| 16 | 0.133649 | True |
| 17 | 0.000052 | False |
| 18 | 0.079480 | False |
| 19 | 0.111152 | True |
| 20 | 0.025581 | True |
| 21 | 0.013207 | False |
| 22 | 0.102643 | True |
| 23 | 0.014241 | True |
| 24 | 0.093412 | True |
| 25 | 0.000461 | False |
| 26 | 0.206803 | True |
| 27 | 0.004577 | True |
| 28 | 0.173810 | True |
| 29 | 0.189945 | False |
| ... | ... | ... |
| 2045 | 0.000076 | True |
| 2046 | 0.044173 | False |
| 2047 | 0.086346 | True |
| 2048 | 0.032253 | False |

| | | |
|------|----------|-------|
| 2049 | 0.029705 | True |
| 2050 | 0.134203 | True |
| 2051 | 0.038915 | True |
| 2052 | 0.026321 | False |
| 2053 | 0.026236 | True |
| 2054 | 0.000903 | True |
| 2055 | 0.038915 | True |
| 2056 | 0.111411 | True |
| 2057 | 0.073482 | True |
| 2058 | 0.105506 | True |
| 2059 | 0.027351 | True |
| 2060 | 0.007958 | True |
| 2061 | 0.001498 | True |
| 2062 | 0.004633 | True |
| 2063 | 0.082981 | False |
| 2064 | 0.008167 | True |
| 2065 | 0.029248 | True |
| 2066 | 0.118184 | True |
| 2067 | 0.076507 | True |
| 2068 | 0.193548 | False |
| 2069 | 0.076902 | False |
| 2070 | 0.175219 | True |
| 2071 | 0.040836 | False |
| 2072 | 0.031379 | True |
| 2073 | 0.068957 | True |
| 2074 | 0.076110 | False |

[2075 rows x 12 columns]

```
In [19]: api_df_now.head()
```

```
Out[19]:
```

| | favorite_count | retweet_count | tweet_id |
|---|----------------|---------------|--------------------|
| 0 | 39467 | 8853 | 892420643555336193 |
| 1 | 33819 | 6514 | 892177421306343426 |
| 2 | 25461 | 4328 | 891815181378084864 |
| 3 | 42908 | 8964 | 891689557279858688 |
| 4 | 41048 | 9774 | 891327558926688256 |

```
In [20]: api_df_now.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2354 entries, 0 to 2353
Data columns (total 3 columns):
favorite_count    2354 non-null int64
retweet_count     2354 non-null int64
tweet_id         2354 non-null int64
dtypes: int64(3)
memory usage: 55.2 KB
```

```
In [21]: weratedogs_df.rating_denominator.value_counts()
```

```
Out[21]: 10      2333
         11        3
         50        3
         80        2
         20        2
          2         1
         16        1
         40        1
         70        1
         15        1
         90        1
        110        1
        120        1
        130        1
        150        1
        170        1
          7         1
          0         1
         Name: rating_denominator, dtype: int64
```

```
In [22]: weratedogs_df.rating_numerator.value_counts()
```

```
Out[22]: 12      558
         11     464
         10     461
         13     351
          9     158
          8     102
          7      55
         14      54
          5      37
          6      32
          3      19
          4      17
          1       9
          2       9
        420       2
          0       2
         15       2
         75       2
         80       1
         20       1
         24       1
         26       1
         44       1
         50       1
```

```

60      1
165     1
84      1
88      1
144     1
182     1
143     1
666     1
960     1
1776    1
17      1
27      1
45      1
99      1
121     1
204     1
Name: rating_numerator, dtype: int64

```

```

In [23]: weratedogs_clean_df = weratedogs_df.copy()
        photos_clean_df = predictions_df.copy()
        api_clean_df = api_df_now.copy()

```

```

In [24]: weratedogs_clean_df.to_csv('weratedogs_clean.csv',encoding='utf-8',index=False)
        photos_clean_df.to_csv('predictions_clean.csv',encoding='utf-8',index=False)
        api_clean_df.to_csv('api_clean.csv',encoding='utf-8',index=False)

```

```

In [25]: weratedogs_clean_df = pd.read_csv('weratedogs_clean.csv')
        predictions_clean_df = pd.read_csv('predictions_clean.csv')
        api_clean_df = pd.read_csv('api_clean.csv')

```

```

In [26]: first_archive_master_df = pd.merge(weratedogs_clean_df, api_clean_df, on=['tweet_id','t

```

```

In [27]: first_archive_master_df.head(2)

```

```

Out[27]:
      tweet_id  in_reply_to_status_id  in_reply_to_user_id  \
0  892420643555336193                NaN                NaN
1  892177421306343426                NaN                NaN

      timestamp  \
0  2017-08-01 16:23:56 +0000
1  2017-08-01 00:17:27 +0000

      source  \
0  <a href="http://twitter.com/download/iphone" r...
1  <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

```

```

retweeted_status_user_id retweeted_status_timestamp \
0      NaN      NaN
1      NaN      NaN

expanded_urls rating_numerator \
0 https://twitter.com/dog_rates/status/892420643... 13
1 https://twitter.com/dog_rates/status/892177421... 13

rating_denominator name doggo floofer pupper puppo favorite_count \
0      10 Phineas None None None None 39467.0
1      10 Tilly None None None None 33819.0

retweet_count
0      8853.0
1      6514.0

```

In [28]: first_archive_master_2_df = pd.merge(first_archive_master_df, predictions_clean_df, on=

In [29]: first_archive_master_2_df.head(2)

```

Out[29]:
tweet_id in_reply_to_status_id in_reply_to_user_id \
0 892420643555336193      NaN      NaN
1 892177421306343426      NaN      NaN

timestamp \
0 2017-08-01 16:23:56 +0000
1 2017-08-01 00:17:27 +0000

source \
0 <a href="http://twitter.com/download/iphone" r...
1 <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

retweeted_status_user_id retweeted_status_timestamp \
0      NaN      NaN
1      NaN      NaN

expanded_urls ... img_num \
0 https://twitter.com/dog_rates/status/892420643... ... 1.0
1 https://twitter.com/dog_rates/status/892177421... ... 1.0

p1 p1_conf p1_dog p2 p2_conf p2_dog p3 p3_conf \
0 orange 0.097049 False bagel 0.085851 False banana 0.076110
1 Chihuahua 0.323581 True Pekinese 0.090647 True papillon 0.068957

```

```

      p3_dog
0  False
1   True

[2 rows x 30 columns]

```

```
In [30]: first_archive_master_df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2356 entries, 0 to 2355
Data columns (total 19 columns):
tweet_id                2356 non-null int64
in_reply_to_status_id   78 non-null float64
in_reply_to_user_id     78 non-null float64
timestamp               2356 non-null object
source                  2356 non-null object
text                    2356 non-null object
retweeted_status_id      181 non-null float64
retweeted_status_user_id 181 non-null float64
retweeted_status_timestamp 181 non-null object
expanded_urls           2297 non-null object
rating_numerator         2356 non-null int64
rating_denominator       2356 non-null int64
name                     2356 non-null object
doggo                    2356 non-null object
floofer                  2356 non-null object
pupper                  2356 non-null object
puppo                    2356 non-null object
favorite_count           2354 non-null float64
retweet_count            2354 non-null float64
dtypes: float64(6), int64(3), object(10)
memory usage: 368.1+ KB

```

```
In [31]: first_archive_master_2_df.to_csv('first_archive_master_2.csv',encoding='utf-8',index=False)
```

```
In [32]: df = pd.read_csv('first_archive_master_2.csv')
```

```
In [33]: df = df[df.retweeted_status_id.isnull()]
```

```
In [34]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2175 entries, 0 to 2355
Data columns (total 30 columns):
tweet_id                2175 non-null int64
in_reply_to_status_id   78 non-null float64
in_reply_to_user_id     78 non-null float64

```

```

timestamp                2175 non-null object
source                  2175 non-null object
text                    2175 non-null object
retweeted_status_id      0 non-null float64
retweeted_status_user_id 0 non-null float64
retweeted_status_timestamp 0 non-null object
expanded_urls            2117 non-null object
rating_numerator         2175 non-null int64
rating_denominator       2175 non-null int64
name                    2175 non-null object
doggo                   2175 non-null object
floofer                 2175 non-null object
pupper                  2175 non-null object
puppo                   2175 non-null object
favorite_count           2175 non-null float64
retweet_count            2175 non-null float64
jpg_url                  1994 non-null object
img_num                  1994 non-null float64
p1                       1994 non-null object
p1_conf                  1994 non-null float64
p1_dog                   1994 non-null object
p2                       1994 non-null object
p2_conf                  1994 non-null float64
p2_dog                   1994 non-null object
p3                       1994 non-null object
p3_conf                  1994 non-null float64
p3_dog                   1994 non-null object
dtypes: float64(10), int64(3), object(17)
memory usage: 526.8+ KB

```

```
In [35]: df = df.drop(['retweeted_status_id', 'retweeted_status_user_id', 'retweeted_status_timestamp'])
```

```
In [36]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2175 entries, 0 to 2355
Data columns (total 27 columns):
tweet_id                2175 non-null int64
in_reply_to_status_id    78 non-null float64
in_reply_to_user_id      78 non-null float64
timestamp                2175 non-null object
source                  2175 non-null object
text                    2175 non-null object
expanded_urls            2117 non-null object
rating_numerator         2175 non-null int64
rating_denominator       2175 non-null int64
name                    2175 non-null object

```

```

doggo                2175 non-null object
floofer              2175 non-null object
pupper               2175 non-null object
puppo                2175 non-null object
favorite_count       2175 non-null float64
retweet_count        2175 non-null float64
jpg_url              1994 non-null object
img_num              1994 non-null float64
p1                   1994 non-null object
p1_conf              1994 non-null float64
p1_dog               1994 non-null object
p2                   1994 non-null object
p2_conf              1994 non-null float64
p2_dog               1994 non-null object
p3                   1994 non-null object
p3_conf              1994 non-null float64
p3_dog               1994 non-null object
dtypes: float64(8), int64(3), object(16)
memory usage: 475.8+ KB

```

```
In [37]: df.head(2)
```

```

Out[37]:
   tweet_id  in_reply_to_status_id  in_reply_to_user_id \
0  892420643555336193             NaN                 NaN
1  892177421306343426             NaN                 NaN

   timestamp \
0  2017-08-01 16:23:56 +0000
1  2017-08-01 00:17:27 +0000

   source \
0  <a href="http://twitter.com/download/iphone" r...
1  <a href="http://twitter.com/download/iphone" r...

   text \
0  This is Phineas. He's a mystical boy. Only eve...
1  This is Tilly. She's just checking pup on you...

   expanded_urls  rating_numerator \
0  https://twitter.com/dog_rates/status/892420643...  13
1  https://twitter.com/dog_rates/status/892177421...  13

   rating_denominator  name  ...  img_num  p1  p1_conf  p1_dog \
0                   10  Phineas  ...    1.0  orange  0.097049  False
1                   10   Tilly  ...    1.0  Chihuahua  0.323581   True

   p2  p2_conf  p2_dog  p3  p3_conf  p3_dog

```


| | | | | | | |
|---|----------|----------|-------|----------|----------|-------|
| 0 | bagel | 0.085851 | False | banana | 0.076110 | False |
| 1 | Pekinese | 0.090647 | True | papillon | 0.068957 | True |

[2 rows x 27 columns]

```
In [38]: df = df[df.in_reply_to_status_id.isnull()]
```

```
In [39]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 27 columns):
tweet_id                2097 non-null int64
in_reply_to_status_id   0 non-null float64
in_reply_to_user_id     0 non-null float64
timestamp               2097 non-null object
source                  2097 non-null object
text                    2097 non-null object
expanded_urls           2094 non-null object
rating_numerator         2097 non-null int64
rating_denominator       2097 non-null int64
name                    2097 non-null object
doggo                   2097 non-null object
floofer                 2097 non-null object
pupper                  2097 non-null object
puppo                   2097 non-null object
favorite_count           2097 non-null float64
retweet_count            2097 non-null float64
jpg_url                 1971 non-null object
img_num                 1971 non-null float64
p1                       1971 non-null object
p1_conf                 1971 non-null float64
p1_dog                  1971 non-null object
p2                       1971 non-null object
p2_conf                 1971 non-null float64
p2_dog                  1971 non-null object
p3                       1971 non-null object
p3_conf                 1971 non-null float64
p3_dog                  1971 non-null object
dtypes: float64(8), int64(3), object(16)
memory usage: 458.7+ KB
```

```
In [40]: df = df.drop(['in_reply_to_status_id', 'in_reply_to_user_id'], axis=1)
```

```
In [41]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
```

Data columns (total 25 columns):

| | |
|--------------------|-----------------------|
| tweet_id | 2097 non-null int64 |
| timestamp | 2097 non-null object |
| source | 2097 non-null object |
| text | 2097 non-null object |
| expanded_urls | 2094 non-null object |
| rating_numerator | 2097 non-null int64 |
| rating_denominator | 2097 non-null int64 |
| name | 2097 non-null object |
| doggo | 2097 non-null object |
| floofer | 2097 non-null object |
| pupper | 2097 non-null object |
| puppo | 2097 non-null object |
| favorite_count | 2097 non-null float64 |
| retweet_count | 2097 non-null float64 |
| jpg_url | 1971 non-null object |
| img_num | 1971 non-null float64 |
| p1 | 1971 non-null object |
| p1_conf | 1971 non-null float64 |
| p1_dog | 1971 non-null object |
| p2 | 1971 non-null object |
| p2_conf | 1971 non-null float64 |
| p2_dog | 1971 non-null object |
| p3 | 1971 non-null object |
| p3_conf | 1971 non-null float64 |
| p3_dog | 1971 non-null object |

dtypes: float64(6), int64(3), object(16)
memory usage: 426.0+ KB

In [42]: df.head(2)

```
Out[42]:
```

| | tweet_id | timestamp | source | text | expanded_urls | rating_numerator |
|---|--------------------|---------------------------|---|---|---|------------------|
| 0 | 892420643555336193 | 2017-08-01 16:23:56 +0000 | <a href="http://twitter.com/download/iphone" r... | This is Phineas. He's a mystical boy. Only eve... | https://twitter.com/dog_rates/status/892420643... | 13 |
| 1 | 892177421306343426 | 2017-08-01 00:17:27 +0000 | <a href="http://twitter.com/download/iphone" r... | This is Tilly. She's just checking pup on you... | https://twitter.com/dog_rates/status/892177421... | 13 |

| | rating_denominator | name | doggo | floofer | ... | img_num | p1 | \ |
|---|--------------------|---------|-------|---------|-----|---------|-----------|---|
| 0 | 10 | Phineas | None | None | ... | 1.0 | orange | |
| 1 | 10 | Tilly | None | None | ... | 1.0 | Chihuahua | |

| | p1_conf | p1_dog | p2 | p2_conf | p2_dog | p3 | p3_conf | p3_dog |
|---|----------|--------|----------|----------|--------|----------|----------|--------|
| 0 | 0.097049 | False | bagel | 0.085851 | False | banana | 0.076110 | False |
| 1 | 0.323581 | True | Pekinese | 0.090647 | True | papillon | 0.068957 | True |

[2 rows x 25 columns]

```
In [43]: df = df[df.jpg_url.notnull()]
```

```
In [44]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1971 entries, 0 to 2355
Data columns (total 25 columns):
tweet_id          1971 non-null int64
timestamp         1971 non-null object
source            1971 non-null object
text              1971 non-null object
expanded_urls     1971 non-null object
rating_numerator  1971 non-null int64
rating_denominator 1971 non-null int64
name              1971 non-null object
doggo             1971 non-null object
floofer           1971 non-null object
pupper           1971 non-null object
puppo            1971 non-null object
favorite_count    1971 non-null float64
retweet_count     1971 non-null float64
jpg_url           1971 non-null object
img_num           1971 non-null float64
p1               1971 non-null object
p1_conf           1971 non-null float64
p1_dog           1971 non-null object
p2               1971 non-null object
p2_conf           1971 non-null float64
p2_dog           1971 non-null object
p3               1971 non-null object
p3_conf           1971 non-null float64
p3_dog           1971 non-null object
dtypes: float64(6), int64(3), object(16)
memory usage: 400.4+ KB
```

```
In [45]: df['doggo'].value_counts()
```

```
Out[45]: None      1898
         doggo      73
         Name: doggo, dtype: int64
```

```
In [46]: df['floofer'].value_counts()
```

```
Out[46]: None      1963
         floofer      8
         Name: floofer, dtype: int64
```

```
In [47]: df['pupper'].value_counts()
```

```
Out[47]: None      1762
         pupper     209
         Name: pupper, dtype: int64
```

```
In [48]: df['puppo'].value_counts()
```

```
Out[48]: None      1948
         puppo      23
         Name: puppo, dtype: int64
```

```
In [49]: df.columns
```

```
Out[49]: Index(['tweet_id', 'timestamp', 'source', 'text', 'expanded_urls',
               'rating_numerator', 'rating_denominator', 'name', 'doggo', 'floofer',
               'pupper', 'puppo', 'favorite_count', 'retweet_count', 'jpg_url',
               'img_num', 'p1', 'p1_conf', 'p1_dog', 'p2', 'p2_conf', 'p2_dog', 'p3',
               'p3_conf', 'p3_dog'],
              dtype='object')
```

```
In [50]: df1 = pd.melt(df, id_vars = ['tweet_id', 'timestamp', 'source', 'text', 'expanded_urls',
               'rating_numerator', 'rating_denominator', 'name', 'favorites', 'retweet_count',
               'p1', 'p1_conf', 'p1_dog', 'p2', 'p2_conf', 'p2_dog', 'p3', 'p3_conf',
               'p3_dog'],
               value_vars = ['doggo', 'floofer', 'pupper', 'puppo'],
               var_name = 'types', value_name = 'dog_type')
```

/opt/conda/lib/python3.6/site-packages/pandas/core/indexing.py:1472: FutureWarning:
Passing list-likes to .loc or [] with any missing label will raise
KeyError in the future, you can use .reindex() as an alternative.

See the documentation here:

<https://pandas.pydata.org/pandas-docs/stable/indexing.html#deprecate-loc-reindex-listlike>
return self._getitem_tuple(key)

```
In [51]: df1.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7884 entries, 0 to 7883
Data columns (total 23 columns):
tweet_id          7884 non-null int64
timestamp         7884 non-null object
source            7884 non-null object
text              7884 non-null object
expanded_urls     7884 non-null object
rating_numerator  7884 non-null int64
rating_denominator 7884 non-null int64
name              7884 non-null object
favorites         0 non-null float64
retweet_count     7884 non-null float64
jpg_url           7884 non-null object
img_num           7884 non-null float64
p1                7884 non-null object
p1_conf           7884 non-null float64
p1_dog            7884 non-null object
p2                7884 non-null object
p2_conf           7884 non-null float64
p2_dog            7884 non-null object
p3                7884 non-null object
p3_conf           7884 non-null float64
p3_dog            7884 non-null object
types             7884 non-null object
dog_type          7884 non-null object
dtypes: float64(6), int64(3), object(14)
memory usage: 1.4+ MB

```

```
In [52]: df1.dog_type.value_counts()
```

```

Out[52]: None          7571
         pupper        209
         doggo         73
         puppo         23
         floofer         8
         Name: dog_type, dtype: int64

```

```
In [53]: df1 = df1.sort_values('dog_type').drop_duplicates('tweet_id', keep = 'last')
```

```
In [54]: df1.dog_type.value_counts()
```

```

Out[54]: None          1668
         pupper        209
         doggo         63
         puppo         23
         floofer         8
         Name: dog_type, dtype: int64

```

```
In [55]: doggos_df = df.drop(['timestamp', 'source', 'text', 'expanded_urls',
                             'rating_numerator', 'rating_denominator', 'name', 'retweet_count', 'jpg_url', 'i
                             'p1', 'p1_conf', 'p1_dog', 'p2', 'p2_conf', 'p2_dog', 'p3', 'p3_conf',
                             'p3_dog'], axis=1)
```

```
In [56]: pd.set_option("display.max_rows",75)
```

```
In [57]: doggos_df = doggos_df[doggos_df.doggo != 'None']
doggos_df
```

```
Out[57]:
```

| | tweet_id | doggo | floofer | pupper | puppo | favorite_count |
|-----|--------------------|-------|---------|--------|-------|----------------|
| 9 | 890240255349198849 | doggo | None | None | None | 32467.0 |
| 43 | 884162670584377345 | doggo | None | None | None | 20771.0 |
| 99 | 872967104147763200 | doggo | None | None | None | 28031.0 |
| 108 | 871515927908634625 | doggo | None | None | None | 20730.0 |
| 121 | 869596645499047938 | doggo | None | None | None | 16476.0 |
| 172 | 858843525470990336 | doggo | None | None | None | 16304.0 |
| 191 | 855851453814013952 | doggo | None | None | puppo | 47844.0 |
| 200 | 854010172552949760 | doggo | floofer | None | None | 17169.0 |
| 240 | 846514051647705089 | doggo | None | None | None | 48410.0 |
| 248 | 845397057150107648 | doggo | None | None | None | 8241.0 |
| 300 | 836753516572119041 | doggo | None | None | None | 21029.0 |
| 318 | 834574053763584002 | doggo | None | None | None | 14993.0 |
| 331 | 832998151111966721 | doggo | None | None | None | 14549.0 |
| 339 | 832273440279240704 | doggo | None | None | None | 12385.0 |
| 344 | 832032802820481025 | doggo | None | None | None | 13887.0 |
| 345 | 831939777352105988 | doggo | None | None | None | 26404.0 |
| 351 | 831322785565769729 | doggo | None | None | None | 10042.0 |
| 362 | 829449946868879360 | doggo | None | None | None | 11519.0 |
| 363 | 829374341691346946 | doggo | None | None | None | 38074.0 |
| 372 | 828381636999917570 | doggo | None | None | None | 13864.0 |
| 384 | 826958653328592898 | doggo | None | None | None | 23767.0 |
| 385 | 826848821049180160 | doggo | None | None | None | 40325.0 |
| 389 | 826476773533745153 | doggo | None | None | None | 20275.0 |
| 391 | 826204788643753985 | doggo | None | None | None | 5361.0 |
| 423 | 821765923262631936 | doggo | None | None | None | 9317.0 |
| 426 | 821407182352777218 | doggo | None | None | None | 13075.0 |
| 429 | 821107785811234820 | doggo | None | None | None | 10645.0 |
| 440 | 819924195358416896 | doggo | None | None | None | 14305.0 |
| 448 | 819006400881917954 | doggo | None | None | None | 49960.0 |
| 449 | 819004803107983360 | doggo | None | None | None | 95450.0 |
| 460 | 817777686764523521 | doggo | None | pupper | None | 11901.0 |
| 489 | 813910438903693312 | doggo | None | None | None | 10342.0 |
| 493 | 813202720496779264 | doggo | None | None | None | 10192.0 |
| 499 | 813127251579564032 | doggo | None | None | None | 13242.0 |
| 501 | 813096984823349248 | doggo | None | None | None | 11694.0 |
| 531 | 808106460588765185 | doggo | None | pupper | None | 9701.0 |
| 575 | 801115127852503040 | doggo | None | pupper | None | 8992.0 |

| | | | | | | |
|------|--------------------|-------|------|--------|------|----------|
| 585 | 800018252395122689 | doggo | None | None | None | 31768.0 |
| 587 | 799757965289017345 | doggo | None | None | None | 9390.0 |
| 616 | 796484825502875648 | doggo | None | None | None | 8472.0 |
| 624 | 795464331001561088 | doggo | None | None | None | 55683.0 |
| 698 | 786363235746385920 | doggo | None | None | None | 12189.0 |
| 705 | 785639753186217984 | doggo | None | pupper | None | 8735.0 |
| 724 | 782747134529531904 | doggo | None | None | None | 8310.0 |
| 727 | 782305867769217024 | doggo | None | None | None | 18630.0 |
| 780 | 775733305207554048 | doggo | None | None | None | 15413.0 |
| 782 | 775364825476165632 | doggo | None | None | None | 8295.0 |
| 835 | 768193404517830656 | doggo | None | None | None | 12157.0 |
| 839 | 767191397493538821 | doggo | None | None | None | 13643.0 |
| 877 | 760893934457552897 | doggo | None | None | None | 4228.0 |
| 889 | 759793422261743616 | doggo | None | pupper | None | 6620.0 |
| 899 | 758828659922702336 | doggo | None | None | None | 12376.0 |
| 914 | 757393109802180609 | doggo | None | None | None | 6462.0 |
| 919 | 756526248105566208 | doggo | None | None | None | 11506.0 |
| 924 | 755206590534418437 | doggo | None | None | None | 18212.0 |
| 944 | 752682090207055872 | doggo | None | None | None | 6642.0 |
| 945 | 752660715232722944 | doggo | None | None | None | 4878.0 |
| 948 | 752334515931054080 | doggo | None | None | None | 4238.0 |
| 956 | 751583847268179968 | doggo | None | pupper | None | 4849.0 |
| 977 | 750011400160841729 | doggo | None | None | None | 3568.0 |
| 989 | 748932637671223296 | doggo | None | None | None | 6461.0 |
| 992 | 748692773788876800 | doggo | None | None | None | 4659.0 |
| 1030 | 745433870967832576 | doggo | None | None | None | 7824.0 |
| 1039 | 744234799360020481 | doggo | None | None | None | 131075.0 |
| 1063 | 741067306818797568 | doggo | None | pupper | None | 10342.0 |
| 1079 | 739238157791694849 | doggo | None | None | None | 75163.0 |
| 1103 | 735256018284875776 | doggo | None | None | None | 3675.0 |
| 1113 | 733109485275860992 | doggo | None | pupper | None | 44619.0 |
| 1117 | 732375214819057664 | doggo | None | None | None | 8832.0 |
| 1141 | 727644517743104000 | doggo | None | None | None | 6431.0 |
| 1156 | 724771698126512129 | doggo | None | None | None | 2593.0 |
| 1176 | 719991154352222208 | doggo | None | None | None | 5281.0 |
| 1204 | 716080869887381504 | doggo | None | None | None | 5272.0 |

```
In [58]: df1.loc[df1.tweet_id == 817777686764523521, 'dog_type'] = 'doggo'
df1.loc[df1.tweet_id == 808106460588765185, 'dog_type'] = 'None'
df1.loc[df1.tweet_id == 801115127852503040, 'dog_type'] = 'doggo'
df1.loc[df1.tweet_id == 785639753186217984, 'dog_type'] = 'doggo'
df1.loc[df1.tweet_id == 759793422261743616, 'dog_type'] = 'None'
df1.loc[df1.tweet_id == 751583847268179968, 'dog_type'] = 'None'
df1.loc[df1.tweet_id == 741067306818797568, 'dog_type'] = 'None'
df1.loc[df1.tweet_id == 733109485275860992, 'dog_type'] = 'None'
```

```
In [59]: df1 = df1.drop('types',axis=1)
```

```
In [60]: df1.info()
```

```

df1.dog_type.value_counts()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1971 entries, 1895 to 6618
Data columns (total 22 columns):
tweet_id          1971 non-null int64
timestamp         1971 non-null object
source            1971 non-null object
text              1971 non-null object
expanded_urls     1971 non-null object
rating_numerator  1971 non-null int64
rating_denominator 1971 non-null int64
name              1971 non-null object
favorites         0 non-null float64
retweet_count     1971 non-null float64
jpg_url           1971 non-null object
img_num           1971 non-null float64
p1                1971 non-null object
p1_conf           1971 non-null float64
p1_dog            1971 non-null object
p2                1971 non-null object
p2_conf           1971 non-null float64
p2_dog            1971 non-null object
p3                1971 non-null object
p3_conf           1971 non-null float64
p3_dog            1971 non-null object
dog_type          1971 non-null object
dtypes: float64(6), int64(3), object(13)
memory usage: 354.2+ KB

```

```

Out[60]: None          1673
         pupper        201
         doggo         66
         puppo         23
         floofer         8
         Name: dog_type, dtype: int64

```

```
In [61]: df1.head(2)
```

```

Out[61]:
      tweet_id      timestamp \
1895  667405339315146752  2015-11-19 18:13:27 +0000
1894  667435689202614272  2015-11-19 20:14:03 +0000

      source \
1895  <a href="http://twitter.com/download/iphone" r...
1894  <a href="http://twitter.com/download/iphone" r...

      text \

```



```

1895 This is Biden. Biden just tripped... 7/10 http...
1894 Ermergerd 12/10 https://t.co/PQni2sjPsm

```

```

                                expanded_urls  rating_numerator  \
1895 https://twitter.com/dog_rates/status/667405339...          7
1894 https://twitter.com/dog_rates/status/667435689...         12

```

```

        rating_denominator  name  favorites  retweet_count  ...  \
1895                10  Biden          NaN          234.0  ...
1894                10   None          NaN           89.0  ...

```

```

                p1  p1_conf p1_dog                p2  p2_conf p2_dog  \
1895 Saint_Bernard 0.381377   True          Leonberg 0.127998   True
1894   Rottweiler 0.999091   True miniature_pinscher 0.000450   True

```

```

                p3  p3_conf p3_dog  dog_type
1895   golden_retriever 0.069357   True   None
1894 black-and-tan_coonhound 0.000157   True   None

```

```
[2 rows x 22 columns]
```

```

In [62]: breed = []
        confidence = []

```

```

In [63]: def breed_confidence(row):
        if row['p1_dog'] == True:
            breed.append(row['p1'])
            confidence.append(row['p1_conf'])
        elif row['p2_dog'] == True:
            breed.append(row['p2'])
            confidence.append(row['p2_conf'])
        elif row['p3_dog'] == True:
            breed.append(row['p3'])
            confidence.append(row['p3_conf'])
        else:
            breed.append('Unknown')
            confidence.append(0)

```

```

In [64]: df1.apply(breed_confidence, axis=1)

```

```

Out[64]: 1895    None
        1894    None
        1893    None
        1892    None
        1891    None
        1890    None
        1889    None
        1888    None
        1887    None

```

| | |
|------|------|
| 1886 | None |
| 1885 | None |
| 1884 | None |
| 1896 | None |
| 1883 | None |
| 1881 | None |
| 1880 | None |
| 1879 | None |
| 1878 | None |
| 1877 | None |
| 1876 | None |
| 1875 | None |
| 1874 | None |
| 1873 | None |
| 1872 | None |
| 1871 | None |
| 1870 | None |
| 1882 | None |
| 1897 | None |
| 1898 | None |
| 1899 | None |
| 1926 | None |
| 1925 | None |
| 1924 | None |
| 1923 | None |
| 1922 | None |
| 1921 | None |
| 1920 | None |
| | ... |
| 5095 | None |
| 4022 | None |
| 5511 | None |
| 5097 | None |
| 5100 | None |
| 5109 | None |
| 5022 | None |
| 4682 | None |
| 5503 | None |
| 4590 | None |
| 5104 | None |
| 5497 | None |
| 4018 | None |
| 5504 | None |
| 5975 | None |
| 5991 | None |
| 6375 | None |
| 6329 | None |
| 6607 | None |

```

6689    None
6322    None
6207    None
6209    None
6678    None
6219    None
5925    None
5927    None
6018    None
6236    None
6720    None
6059    None
6408    None
6046    None
6390    None
6588    None
6443    None
6618    None
Length: 1971, dtype: object

```

```

In [65]: df1['breed'] = breed
         df1['confidence'] = confidence

```

```

In [66]: df1.head(2)

```

```

Out[66]:
          tweet_id          timestamp \
1895  667405339315146752  2015-11-19 18:13:27 +0000
1894  667435689202614272  2015-11-19 20:14:03 +0000

          source \
1895  <a href="http://twitter.com/download/iphone" r...
1894  <a href="http://twitter.com/download/iphone" r...

          text \
1895  This is Biden. Biden just tripped... 7/10 http...
1894  Ermergerd 12/10 https://t.co/PQni2sjPsm

          expanded_urls  rating_numerator \
1895  https://twitter.com/dog_rates/status/667405339...      7
1894  https://twitter.com/dog_rates/status/667435689...     12

          rating_denominator  name  favorites  retweet_count  ...  p1_dog \
1895              10  Biden      NaN          234.0  ...      True
1894              10   None      NaN           89.0  ...      True

          p2  p2_conf  p2_dog          p3  p3_conf \
1895      Leonberg  0.127998    True      golden_retriever  0.069357
1894  miniature_pinscher  0.000450    True  black-and-tan_coonhound  0.000157

```

| | p3_dog | dog_type | breed | confidence |
|------|--------|----------|---------------|------------|
| 1895 | True | None | Saint_Bernard | 0.381377 |
| 1894 | True | None | Rottweiler | 0.999091 |

[2 rows x 24 columns]

```
In [67]: df1 = df1.drop(['img_num', 'p1', 'p1_conf', 'p1_dog', 'p2', 'p2_conf', 'p2_dog', 'p3', 'p3_conf'])
```

```
In [68]: df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1971 entries, 1895 to 6618
Data columns (total 14 columns):
tweet_id          1971 non-null int64
timestamp         1971 non-null object
source            1971 non-null object
text              1971 non-null object
expanded_urls     1971 non-null object
rating_numerator  1971 non-null int64
rating_denominator 1971 non-null int64
name              1971 non-null object
favorites         0 non-null float64
retweet_count     1971 non-null float64
jpg_url           1971 non-null object
dog_type          1971 non-null object
breed             1971 non-null object
confidence        1971 non-null float64
dtypes: float64(3), int64(3), object(8)
memory usage: 231.0+ KB
```

```
In [69]: list(df1['name'])
```

```
Out[69]: ['Biden',
          'None',
          'None',
          'None',
          'Cupcake',
          'Reese',
          'a',
          'None',
          'Philippe',
          'Hall',
          'None',
          'Dook',
          'Fwed',
          'None',
          'Frank',
```

'a',
'Kial',
'George',
'None',
'None',
'Olive',
'Calvin',
'a',
'None',
'None',
'None',
'Naphaniel',
'None',
'Genevieve',
'Joshwa',
'a',
'None',
'Jockson',
'Jeph',
'None',
'None',
'None',
'a',
'Carll',
'Klevin',
'None',
'Keet',
'None',
'Dave',
'Clybe',
'Gabe',
'None',
'Pipsy',
'Bradlay',
'Churlie',
'Kenneth',
'Clarence',
'None',
'a',
'Timison',
'None',
'None',
'Filup',
'None',
'a',
'Shaggy',
'Josep',
'None',

'Tedrick',
'Stu',
'Erik',
'Cleopatricia',
'None',
'Otis',
'Jiminy',
'None',
'Alfie',
'None',
'a',
'Lugan',
'a',
'Christoper',
'None',
'None',
'None',
'None',
'None',
'None',
'None',
'None',
'None',
'an',
'a',
'an',
'None',
'None',
'None',
'None',
'None',
'a',
'quite',
'Walter',
'None',
'None',
'None',
'None',
'None',
'None',
'Scout',
'Kreggory',
'Jimothy',
'None',
'None',
'None',
'the',
'Phineas',
'a',
'a',

'a',
'None',
'a',
'an',
'a',
'a',
'the',
'None',
'None',
'Cash',
'None',
'Nugget',
'Goose',
'Venti',
'Benedict',
'Nelly',
'Dante',
'Penny',
'Aja',
'Emmy',
'Shadow',
'Beau',
'Derek',
'Mingus',
'Maya',
'None',
'None',
'Canela',
'such',
'Jeffrey',
'Gerald',
'Ralphus',
'Zeke',
'Jim',
'Oliver',
'Ted',
'Bruno',
'Koda',
'Zoey',
'None',
'Jax',
'None',
'Franklin',
'Darla',
'Archie',
'Tilly',
'Waffles',
'Jimbo',

'None',
'Steven',
'Jack',
'Duddles',
'Bailey',
'Romeo',
'None',
'Jesse',
'Bella',
'None',
'Louis',
'Elliot',
'Gary',
'Rey',
'Koko',
'Alfy',
'Stanley',
'Rusty',
'None',
'Bella',
'Noah',
'None',
'Kevin',
'Earl',
'Maisey',
'None',
'Lola',
'None',
'None',
'Lucky',
'Herald',
'Chipson',
'None',
'Carl',
'None',
'None',
'Ricky',
'Clarence',
'Luca',
'Durg',
'None',
'Covach',
'None',
'Geoff',
'Maxwell',
'the',
'Phred',
'Eriq',

'Lolo',
'Fynn',
'Hazel',
'None',
'Brad',
'Jackson',
'None',
'Baxter',
'Samson',
'Gordon',
'Tyrone',
'Kramer',
'None',
'Marq',
'Opal',
'Trooper',
'Louis',
'Kilo',
'None',
'Derby',
'Louis',
'Frönq',
'Pubert',
'None',
'None',
'Cedrick',
'Peaches',
'Sadie',
'Lincoln',
'None',
'Freddery',
'None',
'Kobe',
'Chester',
'Lola',
'None',
'Bodie',
'Phil',
'Archie',
'Brian',
'None',
'Colin',
'Bob',
'None',
'Calvin',
'Wyatt',
'None',
'Olivier',

'Vinscent',
'Dunkin',
'Buddy',
'None',
'None',
'None',
'None',
'Molly',
'None',
'Bailey',
'Leo',
'None',
'Richie',
'None',
'None',
'Brockly',
'Edgar',
'a',
'Cody',
'Amber',
'Chester',
'Tupawc',
'None',
'None',
'Kingsley',
'None',
'Chuck',
'Glacier',
'Asher',
'Carly',
'Apollo',
'None',
'Jimothy',
'None',
'actually',
'Ulysses',
'Tess',
'Samson',
'Jax',
'Todo',
'Brody',
'Kreg',
'None',
'None',
'Joey',
'None',
'Taco',
'None',

'Sarge',
'Champ',
'Ozzy',
'Reggie',
'Goliath',
'Toby',
'CeCe',
'None',
'None',
'Panda',
'Percy',
'Brody',
'Moofasa',
'None',
'Cheesy',
'Griswold',
'Sadie',
'Alice',
'Ozzie',
'None',
'Aspen',
'Penny',
'Darrel',
'Lulu',
'None',
'Bruiser',
'Oreo',
'Jerry',
'Alfie',
'None',
'Rocky',
'None',
'None',
'Perry',
'Olive',
'Lorelei',
'Hammond',
'None',
'Flávio',
'None',
'Crimson',
'Horace',
'Brooks',
'Theodore',
'None',
'None',
'None',
'Jack',

'None',
'Griffin',
'Sebastian',
'Buddy',
'Jiminy',
'Tiger',
'Sadie',
'Lupe',
'None',
'Kulet',
'None',
'Tino',
'None',
'None',
'Obi',
'Jeph',
'Pippa',
'Bobby',
'Berb',
'by',
'None',
'Reptar',
'Karma',
'Vince',
'Cecil',
'Olaf',
'None',
'None',
'Watson',
'Tater',
'Lacy',
'None',
'infuriating',
'Billy',
'Gunner',
'None',
'Sora',
'Staniel',
'Milo',
'None',
'None',
'Gary',
'None',
'None',
'Monkey',
'Beemo',
'Chester',
'Walker',

'None',
'None',
'Clarkus',
'Keurig',
'Harper',
'Jiminus',
'None',
'Sadie',
'Rufus',
'Farfle',
'None',
'Remington',
'Penny',
'Cooper',
'Jebberson',
'River',
'Oliver',
'Bobble',
'Bear',
'Smokey',
'Coco',
'Mister',
'Malikai',
'Klevin',
'Rodney',
'None',
'Olive',
'None',
'Jennifur',
'Link',
'Flávio',
'Ebby',
'a',
'Barclay',
'Suki',
'Alexanderson',
'None',
'Bowie',
'Bluebert',
'None',
'None',
'Otis',
'Sadie',
'Pippa',
'Lucy',
'Sid',
'Harnold',
'Neptune',

'Charleson',
'Carper',
'Steven',
'Stephanus',
'Bubbles',
'Kane',
'Sunny',
'None',
'Stubert',
'Opie',
'Lance',
'None',
'Piper',
'None',
'Kanu',
'None',
'None',
'Thumas',
'None',
'Benedict',
'Curtis',
'Calbert',
'Siba',
'Michelangelo',
'Nico',
'Oscar',
'Bertson',
'Zeus',
'Toby',
'a',
'Travis',
'Cassie',
'Cupcake',
'Kathmandu',
'None',
'Reagan',
'Charl',
'None',
'Calbert',
'Miley',
'None',
'Archie',
'Dave',
'Fillup',
'None',
'Nala',
'Thumas',
'Coops',

'Eazy',
'Dotsy',
'Murphy',
'Kaia',
'Vincent',
'very',
'Daniel',
'Crouton',
'Oliver',
'None',
'None',
'None',
'Rosie',
'Oakley',
'Charlie',
'getting',
'Bentley',
'None',
'Blipson',
'Ole',
'CeCe',
'Cassie',
'Doug',
'None',
'Lily',
'Beau',
'Franklin',
'Maximus',
'Jessiga',
'Claude',
'None',
'Rusty',
'Cuddles',
'Carter',
'Chet',
'None',
'None',
'Reese',
'Terrenth',
'Bode',
'Jimothy',
'Cooper',
'Zara',
'None',
'Max',
'Lucy',
'Chesterson',
'Kara',

'None',
'Koda',
'None',
'Oliver',
'None',
'Shadow',
'Katie',
'Elliot',
'Ellie',
'Tucker',
'Dexter',
'None',
'Lucia',
'Yoda',
'Coopson',
'Bilbo',
'Rilo',
'Fiji',
'None',
'None',
'None',
'a',
'Rambo',
'Lucy',
'Chip',
'an',
'a',
'Socks',
'Charlie',
'Bella',
'None',
'None',
'a',
'Stanley',
'Bisquick',
'Ralphé',
'Emmie',
'Penelope',
'None',
'Amy',
'Jett',
'None',
'Colby',
'None',
'None',
'Toby',
'Phred',
'None',

'Clarence',
'None',
'Remington',
'None',
'None',
'Holly',
'Anthony',
'None',
'Herm',
'None',
'Sparky',
'Lola',
'Lou',
'None',
'Tanner',
'None',
'a',
'Penelope',
'None',
'Maggie',
'None',
'None',
'None',
'Edd',
'None',
'Phineas',
'Bloo',
'Pluto',
'Ester',
'Danny',
'None',
'Jeffrie',
'Sage',
'Jerry',
'Vincent',
'None',
'Traviss',
'None',
'None',
'Creg',
'None',
'Antony',
'Trigger',
'None',
'Mason',
'Andy',
'Gin',
'Paull',

'Malcolm',
'Pip',
'Bradley',
'Cal',
'None',
'Ridley',
'Snickers',
'None',
'None',
'Kreg',
'None',
'None',
'Pippin',
'Buddy',
'light',
'Jackie',
'None',
'Darby',
'None',
'Taz',
'a',
'Caryl',
'Franq',
'Julius',
'just',
'Patch',
'Charlie',
'Hammond',
'Scruffers',
'Alejandro',
'Jax',
'None',
'Reginald',
'Baxter',
'JD',
'None',
'Marley',
'the',
'Tuco',
'Oscar',
'None',
'Mark',
'None',
'space',
'None',
'Dylan',
'Mia',
'Brad',

'Batdog',
'Mojo',
'Lola',
'Gizmo',
'a',
'Koda',
'None',
'Hank',
'Hanz',
'None',
'None',
'Winifred',
'Kevon',
'a',
'Banditt',
'Bobb',
'Amélie',
'Ruffles',
'Lambeau',
'Herald',
'an',
'a',
'Jeremy',
'Chaz',
'None',
'Ruby',
'None',
'Bernie',
'Harrison',
'None',
'Terrance',
'Tucker',
'None',
'None',
'Jaycob',
'Samsom',
'Churlie',
'Timofy',
'Skittles',
'Ron',
'None',
'None',
'Torque',
'Bisquick',
'Jareld',
'None',
'Sophie',
'Rusty',

'None',
'None',
'Zeek',
'a',
'None',
'Jo',
'a',
'Gòrdón',
'Spark',
'None',
'None',
'Marvin',
'a',
'Kallie',
'Jomathan',
'Maks',
'DayZ',
'Jaspers',
'Clarq',
'Ralf',
'None',
'None',
'Ben',
'Tommy',
'Liam',
'Randall',
'Winston',
'Fabio',
'None',
'Wally',
'Sandra',
'a',
'Raphael',
'None',
'a',
'Daisy',
'Larry',
'Nigel',
'Peanut',
'Scooter',
'None',
'Damon',
'Herb',
'a',
'Willy',
'Sam',
'Butters',
'Zoey',

'None',
'Julio',
'None',
'Kirk',
'Sully',
'Dug',
'Saydee',
'None',
'Chip',
'Olivier',
'Billl',
'Ronduh',
'a',
'Kollin',
'Keith',
'Linda',
'None',
'Skye',
'None',
'Penny',
'a',
'Shawwn',
'Louie',
'Kloey',
'Clarence',
'Chester',
'a',
'None',
'Andru',
'Sammy',
'Leonidas',
'Ed',
'None',
'Vinnie',
'Bubbles',
'None',
'the',
'None',
'Penny',
'None',
'Griffin',
'Ash',
'None',
'None',
'Karl',
'Tyrus',
'None',
'Cooper',

'None',
'Chuq',
'Jackson',
'Tassy',
'None',
'Derek',
'Duke',
'None',
'Sammy',
'Ellie',
'a',
'a',
'Lenny',
'Buddah',
'Bernie',
'None',
'None',
'None',
'None',
'None',
'Donny',
'None',
'None',
'Striker',
'Bert',
'None',
'Herm',
'Kenneth',
'Winston',
'None',
'None',
'Humphrey',
'Axel',
'Mia',
'Tug',
'Linda',
'Lola',
'None',
'None',
'Chompsky',
'officially',
'Louis',
'None',
'Rubio',
'Wilson',
'Hurley',
'William',
'a',

'Mike',
'Hunter',
'None',
'Willie',
'Bruce',
'None',
'Evy',
'None',
'Dash',
'None',
'None',
'None',
'Sandy',
'a',
'None',
'None',
'None',
'Sadie',
'Ralph',
'None',
'None',
'Tango',
'None',
'Jessifer',
'Rosie',
'Crumpet',
'None',
'None',
'Jerome',
'Crystal',
'None',
'Grizz',
'Reese',
'Reggie',
'None',
'None',
'Sandy',
'a',
'Kendall',
'Shnuggles',
'George',
'Larry',
'None',
'None',
'Dot',
'None',
'None',
'None',

'a',
'None',
'None',
'None',
'Aiden',
'Acro',
'None',
'Daisy',
'Riley',
'None',
'Steve',
'Oscar',
'None',
'Jimison',
'Terry',
'Gustaf',
'None',
'None',
'Pumpkin',
'Kenzie',
'Dexter',
'None',
'Mac',
'None',
'Frankie',
'None',
'Gerald',
'Bella',
'None',
'Koda',
'None',
'None',
'Lucy',
'None',
'None',
'Maggie',
'None',
'Tedders',
'None',
'Laela',
'Moe',
'a',
'a',
'None',
'Oliver',
'Steven',
'None',
'Coops',


```

'None',
'None',
'None',
'Arnold',
'None',
'Daisy',
'None',
'None',
'None',
'Sophie',
'None',
'Stanley',
'Earl',
'Mitch',
'a',
'Bobbay',
'Henry',
'Bob',
'None',
'Kenny',
'Chesney',
'Gary',
'Lenny',
'Jonah',
'Ava',
...]
```

```
In [70]: faulty_names=[]
```

```
In [71]: def incorrect_name(row):
        if row['name'][0].islower() == True:
            faulty_names.append('None')
        else:
            faulty_names.append(row['name'])
```

```
In [72]: df1.apply(incorrect_name, axis=1)
```

```
Out[72]: 1895    None
         1894    None
         1893    None
         1892    None
         1891    None
         1890    None
         1889    None
         1888    None
         1887    None
         1886    None
         1885    None
         1884    None
```

| | |
|------|------|
| 1896 | None |
| 1883 | None |
| 1881 | None |
| 1880 | None |
| 1879 | None |
| 1878 | None |
| 1877 | None |
| 1876 | None |
| 1875 | None |
| 1874 | None |
| 1873 | None |
| 1872 | None |
| 1871 | None |
| 1870 | None |
| 1882 | None |
| 1897 | None |
| 1898 | None |
| 1899 | None |
| 1926 | None |
| 1925 | None |
| 1924 | None |
| 1923 | None |
| 1922 | None |
| 1921 | None |
| 1920 | None |
| | ... |
| 5095 | None |
| 4022 | None |
| 5511 | None |
| 5097 | None |
| 5100 | None |
| 5109 | None |
| 5022 | None |
| 4682 | None |
| 5503 | None |
| 4590 | None |
| 5104 | None |
| 5497 | None |
| 4018 | None |
| 5504 | None |
| 5975 | None |
| 5991 | None |
| 6375 | None |
| 6329 | None |
| 6607 | None |
| 6689 | None |
| 6322 | None |
| 6207 | None |

```

6209    None
6678    None
6219    None
5925    None
5927    None
6018    None
6236    None
6720    None
6059    None
6408    None
6046    None
6390    None
6588    None
6443    None
6618    None
Length: 1971, dtype: object

```

```

In [73]: df1['faulty_names'] = faulty_names
         df1.head()

```

```

Out[73]:
      tweet_id      timestamp \
1895  667405339315146752  2015-11-19 18:13:27 +0000
1894  667435689202614272  2015-11-19 20:14:03 +0000
1893  667437278097252352  2015-11-19 20:20:22 +0000
1892  667443425659232256  2015-11-19 20:44:47 +0000
1891  667453023279554560  2015-11-19 21:22:56 +0000

      source \
1895  <a href="http://twitter.com/download/iphone" r...
1894  <a href="http://twitter.com/download/iphone" r...
1893  <a href="http://twitter.com/download/iphone" r...
1892  <a href="http://twitter.com/download/iphone" r...
1891  <a href="http://twitter.com" rel="nofollow">Tw...

      text \
1895  This is Biden. Biden just tripped... 7/10 http...
1894  Ermergerd 12/10 https://t.co/PQni2sjPsm
1893  Never seen this breed before. Very pointy pup...
1892  Exotic dog here. Long neck. Weird paws. Obsess...
1891  Meet Cupcake. I would do unspeakable things fo...

      expanded_urls  rating_numerator \
1895  https://twitter.com/dog_rates/status/667405339...      7
1894  https://twitter.com/dog_rates/status/667435689...     12
1893  https://twitter.com/dog_rates/status/667437278...     10
1892  https://twitter.com/dog_rates/status/667443425...      6
1891  https://twitter.com/dog_rates/status/667453023...     11

```

| | rating_denominator | name | favorites | retweet_count | \ |
|------|--------------------|---------|-----------|---------------|---|
| 1895 | 10 | Biden | NaN | 234.0 | |
| 1894 | 10 | None | NaN | 89.0 | |
| 1893 | 10 | None | NaN | 257.0 | |
| 1892 | 10 | None | NaN | 620.0 | |
| 1891 | 10 | Cupcake | NaN | 96.0 | |

| | jpg_url | dog_type | \ |
|------|--|----------|---|
| 1895 | https://pbs.twimg.com/media/CUMZnmhUEAEbtis.jpg | None | |
| 1894 | https://pbs.twimg.com/media/CUM1OHCW4AEgGSi.jpg | None | |
| 1893 | https://pbs.twimg.com/media/CUM2qWaWoAUZO6L.jpg | None | |
| 1892 | https://pbs.twimg.com/media/CUM8QZwW4AAVsBl.jpg | None | |
| 1891 | https://pbs.twimg.com/media/CUNE_OSUwAAAdHhX.jpg | None | |

| | breed | confidence | faulty_names |
|------|--------------------|------------|--------------|
| 1895 | Saint_Bernard | 0.381377 | Biden |
| 1894 | Rottweiler | 0.999091 | None |
| 1893 | Unknown | 0.000000 | None |
| 1892 | Unknown | 0.000000 | None |
| 1891 | Labrador_retriever | 0.825670 | Cupcake |

```
In [74]: df1['faulty_names'].unique()
```

```
Out[74]: array(['Biden', 'None', 'Cupcake', 'Reese', 'Philippe', 'Hall', 'Dook',
                'Fwed', 'Frank', 'Kial', 'George', 'Olive', 'Calvin', 'Naphaniel',
                'Genevieve', 'Joshwa', 'Jockson', 'Jeph', 'Carll', 'Klevin', 'Keet',
                'Dave', 'Clybe', 'Gabe', 'Pipsy', 'Bradlay', 'Churlie', 'Kenneth',
                'Clarence', 'Timison', 'Filup', 'Shaggy', 'Josep', 'Tedrick', 'Stu',
                'Erik', 'Cleopatria', 'Otis', 'Jiminy', 'Alfie', 'Lugan',
                'Christoper', 'Walter', 'Scout', 'Kreggory', 'Jimothy', 'Phineas',
                'Cash', 'Nugget', 'Goose', 'Venti', 'Benedict', 'Nelly', 'Dante',
                'Penny', 'Aja', 'Emmy', 'Shadow', 'Beau', 'Derek', 'Mingus', 'Maya',
                'Canela', 'Jeffrey', 'Gerald', 'Ralphus', 'Zeke', 'Jim', 'Oliver',
                'Ted', 'Bruno', 'Koda', 'Zoey', 'Jax', 'Franklin', 'Darla',
                'Archie', 'Tilly', 'Waffles', 'Jimbo', 'Steven', 'Jack', 'Duddles',
                'Bailey', 'Romeo', 'Jesse', 'Bella', 'Louis', 'Elliot', 'Gary',
                'Rey', 'Koko', 'Alfy', 'Stanley', 'Rusty', 'Noah', 'Kevin', 'Earl',
                'Maisey', 'Lola', 'Lucky', 'Herald', 'Chipson', 'Carl', 'Ricky',
                'Luca', 'Durg', 'Covach', 'Geoff', 'Maxwell', 'Phred', 'Eriq',
                'Lolo', 'Fynn', 'Hazel', 'Brad', 'Jackson', 'Baxter', 'Samson',
                'Gordon', 'Tyrone', 'Kramer', 'Marq', 'Opal', 'Trooper', 'Kilo',
                'Derby', 'Frönq', 'Pubert', 'Cedrick', 'Peaches', 'Sadie',
                'Lincoln', 'Freddery', 'Kobe', 'Chester', 'Bodie', 'Phil', 'Brian',
                'Colin', 'Bob', 'Wyatt', 'Olivier', 'Vinscent', 'Dunkin', 'Buddy',
                'Molly', 'Leo', 'Richie', 'Brockly', 'Edgar', 'Cody', 'Amber',
                'Tupawc', 'Kingsley', 'Chuck', 'Glacier', 'Asher', 'Carly',
                'Apollo', 'Ulysses', 'Tess', 'Todo', 'Brody', 'Kreg', 'Joey',
                'Taco', 'Sarge', 'Champ', 'Ozzy', 'Reggie', 'Goliath', 'Toby',
```

'CeCe', 'Panda', 'Percy', 'Moofasa', 'Cheesy', 'Griswold', 'Alice',
'Ozzie', 'Aspen', 'Darrel', 'Lulu', 'Bruiser', 'Oreo', 'Jerry',
'Rocky', 'Perry', 'Lorelei', 'Hammond', 'Flávio', 'Crimson',
'Horace', 'Brooks', 'Theodore', 'Griffin', 'Sebastian', 'Tiger',
'Lupe', 'Kulet', 'Tino', 'Obi', 'Pippa', 'Bobby', 'Berb', 'Reptar',
'Karma', 'Vince', 'Cecil', 'Olaf', 'Watson', 'Tater', 'Lacy',
'Billy', 'Gunner', 'Sora', 'Staniel', 'Milo', 'Monkey', 'Beemo',
'Walker', 'Clarkus', 'Keurig', 'Harper', 'Jiminus', 'Rufus',
'Farfle', 'Remington', 'Cooper', 'Jebberson', 'River', 'Bobbie',
'Bear', 'Smokey', 'Coco', 'Mister', 'Malikai', 'Rodney', 'Jennifur',
'Link', 'Ebby', 'Barclay', 'Suki', 'Alexanderson', 'Bowie',
'Bluebert', 'Lucy', 'Sid', 'Harnold', 'Neptune', 'Charleson',
'Carper', 'Stephanus', 'Bubbles', 'Kane', 'Sunny', 'Stubert',
'Opie', 'Lance', 'Piper', 'Kanu', 'Thumas', 'Curtis', 'Calbert',
'Siba', 'Michelangelo', 'Nico', 'Oscar', 'Bertson', 'Zeus',
'Travis', 'Cassie', 'Kathmandu', 'Reagan', 'Charl', 'Miley',
'Fillup', 'Nala', 'Coops', 'Eazy', 'Dotsy', 'Murphy', 'Kaia',
'Vincent', 'Daniel', 'Crouton', 'Rosie', 'Oakley', 'Charlie',
'Bentley', 'Blipson', 'Ole', 'Doug', 'Lily', 'Maximus', 'Jessiga',
'Claude', 'Cuddles', 'Carter', 'Chet', 'Terrenth', 'Bode', 'Zara',
'Max', 'Chesterson', 'Kara', 'Katie', 'Ellie', 'Tucker', 'Dexter',
'Lucia', 'Yoda', 'Coopson', 'Bilbo', 'Rilo', 'Fiji', 'Rambo',
'Chip', 'Socks', 'Bisquick', 'Ralphé', 'Emmie', 'Penelope', 'Amy',
'Jett', 'Colby', 'Holly', 'Anthony', 'Herm', 'Sparky', 'Lou',
'Tanner', 'Maggie', 'Edd', 'Bloo', 'Pluto', 'Ester', 'Danny',
'Jeffrie', 'Sage', 'Traviss', 'Creg', 'Antony', 'Trigger', 'Mason',
'Andy', 'Gin', 'Paull', 'Malcolm', 'Pip', 'Bradley', 'Cal',
'Ridley', 'Snickers', 'Pippin', 'Jackie', 'Darby', 'Taz', 'Caryl',
'Franq', 'Julius', 'Patch', 'Scruffers', 'Alejandro', 'Reginald',
'JD', 'Marley', 'Tuco', 'Mark', 'Dylan', 'Mia', 'Batdog', 'Mojo',
'Gizmo', 'Hank', 'Hanz', 'Winifred', 'Kevon', 'Banditt', 'Bobb',
'Amélie', 'Ruffles', 'Lambeau', 'Jeremy', 'Chaz', 'Ruby', 'Bernie',
'Harrison', 'Terrance', 'Jaycob', 'Samsom', 'Timofy', 'Skittles',
'Ron', 'Torque', 'Jareld', 'Sophie', 'Zeek', 'Jo', 'Gördón',
'Spark', 'Marvin', 'Kallie', 'Jomathan', 'Maks', 'DayZ', 'Jaspers',
'Clarq', 'Ralf', 'Ben', 'Tommy', 'Liam', 'Randall', 'Winston',
'Fabio', 'Wally', 'Sandra', 'Raphael', 'Daisy', 'Larry', 'Nigel',
'Peanut', 'Scooter', 'Damon', 'Herb', 'Willy', 'Sam', 'Butters',
'Julio', 'Kirk', 'Sully', 'Dug', 'Saydee', 'Billl', 'Ronduh',
'Kollin', 'Keith', 'Linda', 'Skye', 'Shawwn', 'Louie', 'Kloey',
'Andru', 'Sammy', 'Leonidas', 'Ed', 'Vinnie', 'Ash', 'Karl',
'Tyrus', 'Chuq', 'Juckson', 'Tassy', 'Duke', 'Lenny', 'Buddah',
'Donny', 'Striker', 'Bert', 'Humphrey', 'Axel', 'Tug', 'Chompsky',
'Rubio', 'Wilson', 'Hurley', 'William', 'Mike', 'Hunter', 'Willie',
'Bruce', 'Evy', 'Dash', 'Sandy', 'Ralph', 'Tango', 'Jessifer',
'Crumpet', 'Jerome', 'Crystal', 'Grizz', 'Kendall', 'Shnuggles',
'Dot', 'Aiden', 'Acro', 'Riley', 'Steve', 'Jimison', 'Terry',
'Gustaf', 'Pumpkin', 'Kenzie', 'Mac', 'Frankie', 'Tedders', 'Laela',

'Moe', 'Arnold', 'Mitch', 'Bobbay', 'Henry', 'Kenny', 'Chesney',
'Jonah', 'Ava', 'Rodman', 'Jeb', 'Patrick', 'Roosevelt', 'Bo',
'Django', 'Shooter', 'Sansa', 'Major', 'Lennon', 'Stephan', 'Ito',
'Wallace', 'Kyro', 'Gus', 'Sonny', 'Tyr', 'Pancake', 'Iroh',
'Longfellow', 'Pavlov', 'Chef', 'Mauve', 'Ronnie', 'Severus',
'Winnie', 'Comet', 'Dudley', 'Moose', 'Binky', 'Timber', 'Chloe',
'Tebow', 'Strudel', 'Ken', 'Akumi', 'Eugene', 'Autumn', 'Florence',
'Tom', 'Wafer', 'Finn', 'Anna', 'Jazzy', 'Howie', 'Snicku', 'Titan',
'Augie', 'Halo', 'Mary', 'Brandi', 'Swagger', 'Bauer', 'Baron',
'Layla', 'Eleanor', 'Atlas', 'Sky', 'Oshie', 'Alf', 'Olivia',
'Yogi', 'Timmy', 'Cali', 'Godi', 'Loomis', 'Mairi', 'Stormy',
'Berkeley', 'Hero', 'Doc', 'Dallas', 'Hercules', 'Blue', 'O',
'Solomon', 'Wesley', 'Strider', 'Mya', 'Jay', 'Mattie', 'Dale',
'Monster', 'Robin', 'Nida', 'Benji', 'Newt', 'Maude', 'Juno',
'BeBe', 'Moreton', 'Fiona', 'Dobby', 'Clark', 'Nimbus', 'Mack',
'Remus', 'Levi', 'Logan', 'Rory', 'Butter', 'Stella', 'Rizzy',
'Brownie', 'Eli', 'Happy', 'Mosby', 'Misty', 'Mabel', 'Noosh',
'Kyle', 'Snoop', 'Iggy', 'Sailor', 'General', 'Wiggles', 'Jarod',
'Benny', 'Rumpole', 'Einstein', 'Arya', 'Marlee', 'Cermet', 'Odin',
'Indie', 'Tycho', 'Margo', 'Luna', 'Brady', 'Callie', 'Jarvis',
'Vixen', 'Tuck', 'Furzey', 'Cannon', 'Rontu', 'Georgie', 'Boomer',
'Lili', 'Harold', 'Dewey', 'Clifford', 'Rumble', 'Lassie', 'Kody',
'Harry', 'Meatball', 'Thor', 'Milky', 'Jordy', 'Burt', 'Hobbes',
'Jersey', 'Zooey', 'Quinn', 'Belle', 'Albus', 'Paisley', 'Lorenzo',
'Pawnd', 'Ralphie', 'Boots', 'Brutus', 'Ike', 'Sailer', 'Beebop',
'Mutt', 'Alexander', 'Mo', 'Flash', 'Harlso', 'Crawford', 'Sweet',
'Mookie', 'Arlo', 'Odie', 'Sojourner', 'Monty', 'Maddie', 'Dawn',
'Lilly', 'Tobi', 'Dutch', 'Kuyu', 'Eevee', 'Willow', 'Poppy',
'Bronte', 'Lipton', 'Sunshine', 'Orion', 'Dex', 'Charles',
'Grizzie', 'Chase', 'Tayzie', 'Ace', 'Nollie', 'Rorie', 'Simba',
'Gert', 'Emma', 'Kirby', 'Chadrick', 'Fido', 'Rocco', 'Theo',
'Rose', 'Albert', 'Spencer', 'Storkson', 'Devón', 'Remy', 'Beckham',
'Leonard', 'Corey', 'Brandonald', 'Boston', 'Lilli', 'Bonaparte',
'Bookstore', 'Huxley', 'Stark', 'Beya', 'Cora', 'Tripp', 'Atticus',
'Sugar', 'Arlen', 'Gustav', 'Shiloh', 'Abby', 'Harvey', 'Blanket',
'Geno', 'Stewie', 'Arnie', 'Meyer', 'Gilbert', 'Lilah', 'Spanky',
'Dietrich', 'Kayla', 'Maxaroni', 'Bell', 'Jameson', 'Pupcasso',
'Philbert', 'Bloop', 'Flurpson', 'Blitz', 'Puff', 'Tito', 'Rupert',
'Brudge', 'Ziva', 'Ollie', 'Shadoe', 'Stefan', 'Fred', 'Coleman',
'Chevy', 'Carbon', 'Dixie', 'Klein', 'Cilantro', 'Al', 'Fizz',
'DonDon', 'Sprout', 'Karll', 'Enchilada', 'Raymond', 'Davey',
'Dakota', 'Grey', 'Schnitzel', 'Livvie', 'Glenn', 'Heinrich',
'Loki', 'Ivar', 'Shakespeare', 'Jangle', 'Finley', 'Sephie',
'Sprinkles', 'Shelby', 'Hermione', 'Leela', 'Ralpher', 'Bungalo',
'Gromit', 'Tove', 'Aubie', 'Kota', 'Aldrick', 'Brat', 'Angel',
'Doobert', 'Meera', 'Wishes', 'Astrid', 'Barney', 'Pete', 'Emanuel',
'Miguel', 'Rhino', 'Smiley', 'Mimosa', 'Napolean', 'Rizzo',
'Sampson', 'Deacon', 'Cupid', 'Divine', 'Chubbs', 'Sundance',

```
'Pilot', 'Lenox', 'Anakin', 'Sobe', 'Duchess', 'Kellogg', 'Petrick',
'Blu', 'Grizzwald', 'Clyde', 'Baloo', 'Ralphson', 'Rooney', 'Aqua',
'Edmund', 'Roscoe', 'Buckley', 'Ashleigh', 'Lillie', 'Luther',
'Adele', 'Finnegus', 'Rueben', 'Chuckles', 'Ambrose', 'Trip',
'Dido', 'Huck', 'Godzilla', 'Mona', 'Birf', 'Blakely', 'Kona',
'Millie', 'Hamrick', 'Pepper', 'Willem', 'Lizzie', 'Chelsea',
'Hector', 'Banjo', 'Laika', 'Toffee', 'Brandy', 'Bones', 'Kawhi',
'Marty', 'Hubertson', 'Craig', 'Gerbald', 'Dwight', 'Rinna',
'Pinot', 'Cheryl', 'Trevith', 'Schnozz', 'Zuzu', 'Eve', 'Fletcher',
'Norman', 'Scott', 'Rover', 'Jamesy', 'Zoe', 'Bubba', 'Rolf',
'Jazz', 'Jeffri', 'Grady', 'Ginger', 'Herschel', 'Rufio', 'Gidget',
'Superpup', 'Mollie', 'Pickles', 'Sierra', 'Kaiya', 'Jed', 'Obie',
'Snoopy', 'Diogi', 'Pablo', 'Stuart', 'Shikha', 'Bayley', 'Tonks',
'Betty'], dtype=object)
```

```
In [75]: df1['name'] = df1['faulty_names']
```

```
In [76]: df1 = df1.drop('faulty_names', axis=1)
```

```
In [77]: df1['name'].value_counts()
```

```
Out[77]: None          622
Charlie          11
Lucy             10
Cooper           10
Oliver           10
Penny            9
Tucker           9
Winston          8
Sadie            8
Toby             7
Daisy            7
Lola             7
Jax              6
Stanley          6
Koda             6
Bo              6
Bella           6
Milo             5
Buddy           5
Leo             5
Louis           5
Rusty           5
Chester         5
Dave            5
Scout           5
Bailey          5
Oscar           5
Gary            4
```

| | |
|-----------|---|
| Alfie | 4 |
| Larry | 4 |
| Phil | 4 |
| Bear | 4 |
| Jerry | 4 |
| Dexter | 4 |
| Duke | 4 |
| Reggie | 4 |
| Clarence | 4 |
| ... | |
| Tuck | 1 |
| Pinot | 1 |
| Ziva | 1 |
| Clybe | 1 |
| Gidget | 1 |
| Banjo | 1 |
| Harper | 1 |
| Champ | 1 |
| Monty | 1 |
| Josep | 1 |
| Walker | 1 |
| Cheesy | 1 |
| Burt | 1 |
| Tater | 1 |
| Tassy | 1 |
| Lilli | 1 |
| Jersey | 1 |
| Butter | 1 |
| Clifford | 1 |
| Ambrose | 1 |
| Fwed | 1 |
| Amy | 1 |
| Willie | 1 |
| Alexander | 1 |
| Gordon | 1 |
| Todo | 1 |
| Carper | 1 |
| Akumi | 1 |
| Ronduh | 1 |
| Canela | 1 |
| Darby | 1 |
| Huck | 1 |
| Vinscent | 1 |
| Duddles | 1 |
| Alfy | 1 |
| Luther | 1 |
| Steve | 1 |

Name: name, Length: 913, dtype: int64


```
In [78]: df1.rating_denominator.value_counts()
```

```
Out[78]: 10      1954
         50       3
         80       2
         11       2
        170       1
        150       1
        120       1
        110       1
         90       1
         70       1
         40       1
         20       1
          7       1
          2       1
        Name: rating_denominator, dtype: int64
```

```
In [79]: df1.columns
```

```
Out[79]: Index(['tweet_id', 'timestamp', 'source', 'text', 'expanded_urls',
               'rating_numerator', 'rating_denominator', 'name', 'favorites',
               'retweet_count', 'jpg_url', 'dog_type', 'breed', 'confidence'],
              dtype='object')
```

```
In [80]: rating_df = df1.drop(['timestamp', 'source', 'expanded_urls', 'name', 'favorites', 'retweet_
               'jpg_url', 'dog_type', 'breed', 'confidence'], axis=1)
```

```
In [81]: rating_df = rating_df[rating_df.rating_denominator != 10]
```

```
In [82]: rating_df
```

```
Out[82]:
```

| | tweet_id | text \ |
|------|--------------------|---|
| 1950 | 666287406224695296 | This is an Albanian 3 1/2 legged Episcopalian... |
| 3284 | 682962037429899265 | This is Darrel. He just robbed a 7/11 and is i... |
| 3259 | 684222868335505415 | Someone help the girl is being mugged. Several... |
| 2934 | 710658690886586372 | Here's a brigade of puppies. All look very pre... |
| 2952 | 709198395643068416 | From left to right:\nCletus, Jerome, Alejandro... |
| 2889 | 716439118184652801 | This is Bluebert. He just saw that both #Final... |
| 2910 | 713900603437621249 | Happy Saturday here's 9 puppies on a bench. 99... |
| 3091 | 697463031882764288 | Happy Wednesday here's a bucket of pups. 44/40... |
| 3016 | 704054845121142784 | Here is a whole flock of puppies. 60/50 I'll ... |
| 3449 | 675853064436391936 | Here we have an entire platoon of puppies. Tot... |
| 3391 | 677716515794329600 | IT'S PUPPERGEDDON. Total of 144/120 ...I think... |
| 2353 | 810984652412424192 | Meet Sam. She smiles 24/7 & secretly aspir... |
| 2291 | 820690176645140481 | The floofs have been released I repeat the flo... |
| 2765 | 740373189193256964 | After so many requests, this is Bretagne. She ... |
| 2629 | 758467244762497024 | Why does this never happen at my front door... |
| 2855 | 722974582966214656 | Happy 4/20 from the squad! 13/10 for all https... |

```
2813 731156023742988288 Say hello to this unbelievably well behaved sq...
```

| | rating_numerator | rating_denominator |
|------|------------------|--------------------|
| 1950 | 1 | 2 |
| 3284 | 7 | 11 |
| 3259 | 121 | 110 |
| 2934 | 80 | 80 |
| 2952 | 45 | 50 |
| 2889 | 50 | 50 |
| 2910 | 99 | 90 |
| 3091 | 44 | 40 |
| 3016 | 60 | 50 |
| 3449 | 88 | 80 |
| 3391 | 144 | 120 |
| 2353 | 24 | 7 |
| 2291 | 84 | 70 |
| 2765 | 9 | 11 |
| 2629 | 165 | 150 |
| 2855 | 4 | 20 |
| 2813 | 204 | 170 |

```
In [83]: df1.loc[df1.tweet_id == 666287406224695296, 'rating_numerator'] = 9
         df1.loc[df1.tweet_id == 666287406224695296, 'rating_denominator'] = 10
```

```
In [84]: df1.loc[df1.tweet_id == 716439118184652801, 'rating_numerator'] = 11
         df1.loc[df1.tweet_id == 716439118184652801, 'rating_denominator'] = 10
```

```
In [85]: df1.loc[df1.tweet_id == 810984652412424192, 'rating_numerator'] = 10
         df1.loc[df1.tweet_id == 810984652412424192, 'rating_denominator'] = 10
```

```
In [86]: df1.loc[df1.tweet_id == 722974582966214656, 'rating_numerator'] = 13
         df1.loc[df1.tweet_id == 722974582966214656, 'rating_denominator'] = 10
```

```
In [87]: df1.loc[df1.tweet_id == 740373189193256964, 'rating_numerator'] = 14
         df1.loc[df1.tweet_id == 740373189193256964, 'rating_denominator'] = 10
```

```
In [88]: df1.rating_denominator.value_counts()
```

```
Out[88]: 10      1959
         80       2
         50       2
         170      1
         150      1
         120      1
         110      1
          90      1
          70      1
          40      1
          11      1
         Name: rating_denominator, dtype: int64
```

```
In [89]: df1.rating_numerator.value_counts()
```

```
Out[89]: 12      446
         10      418
         11      393
         13      254
          9      150
          8       95
          7       52
         14       34
          5       33
          6       32
          3       19
          4       15
          2        9
          1        4
        1776        1
         26        1
        420        1
         27        1
         44        1
         45        1
         60        1
         75        1
         80        1
         84        1
         88        1
         99        1
        121        1
        144        1
        165        1
        204        1
          0         1
        Name: rating_numerator, dtype: int64
```

```
In [93]: df1['source'] = df1['source'].apply(lambda x: re.findall(r'>(.*?)<', x)[0])
```

```
In [94]: df1['source'].value_counts()
```

```
Out[94]: Twitter for iPhone      1932
         Twitter Web Client       28
         TweetDeck               11
        Name: source, dtype: int64
```

```
In [95]: df1.timestamp = df1.timestamp.str.strip('+0000')
```

```
In [96]: df1.timestamp = pd.to_datetime(df1.timestamp)
```

```
In [97]: df1.info()
         df1.head(2)
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1971 entries, 1895 to 6618
Data columns (total 14 columns):
tweet_id          1971 non-null int64
timestamp         1971 non-null datetime64[ns]
source            1971 non-null object
text              1971 non-null object
expanded_urls     1971 non-null object
rating_numerator  1971 non-null int64
rating_denominator 1971 non-null int64
name              1971 non-null object
favorites         0 non-null float64
retweet_count     1971 non-null float64
jpg_url           1971 non-null object
dog_type          1971 non-null object
breed             1971 non-null object
confidence        1971 non-null float64
dtypes: datetime64[ns](1), float64(3), int64(3), object(7)
memory usage: 231.0+ KB

```

```

Out[97]:
      tweet_id      timestamp      source \
1895  667405339315146752  2015-11-19 18:13:27  Twitter for iPhone
1894  667435689202614272  2015-11-19 20:14:03  Twitter for iPhone

      text \
1895  This is Biden. Biden just tripped... 7/10 http...
1894  Ermergerd 12/10 https://t.co/PQni2sjPsm

      expanded_urls  rating_numerator \
1895  https://twitter.com/dog_rates/status/667405339...      7
1894  https://twitter.com/dog_rates/status/667435689...     12

      rating_denominator  name  favorites  retweet_count \
1895          10  Biden      NaN      234.0
1894          10   None      NaN      89.0

      jpg_url  dog_type      breed \
1895  https://pbs.twimg.com/media/CUMZnmhUEAEBtis.jpg    None  Saint_Bernard
1894  https://pbs.twimg.com/media/CUM1OHCW4AEgGSi.jpg    None    Rottweiler

      confidence
1895    0.381377
1894    0.999091

```

```

In [99]: df1['tweet_id'] = df1['tweet_id'].astype(object)
df1['source'] = df1['source'].astype('category')
df1['retweet_count'] = df1['retweet_count'].astype(int)

```

```
df1['dog_type'] = df1['dog_type'].astype('category')
df1['confidence'] = df1['confidence'].round(2)
```

```
In [100]: df1.info()
          df1.head(2)
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1971 entries, 1895 to 6618
Data columns (total 14 columns):
tweet_id      1971 non-null object
timestamp     1971 non-null datetime64[ns]
source        1971 non-null category
text          1971 non-null object
expanded_urls  1971 non-null object
rating_numerator  1971 non-null int64
rating_denominator  1971 non-null int64
name          1971 non-null object
favorites      0 non-null float64
retweet_count  1971 non-null int64
jpg_url       1971 non-null object
dog_type      1971 non-null category
breed         1971 non-null object
confidence     1971 non-null float64
dtypes: category(2), datetime64[ns](1), float64(2), int64(3), object(6)
memory usage: 204.3+ KB
```

```
Out[100]:
```

| | tweet_id | timestamp | source | |
|------|--------------------|---------------------|--------------------|--|
| 1895 | 667405339315146752 | 2015-11-19 18:13:27 | Twitter for iPhone | |
| 1894 | 667435689202614272 | 2015-11-19 20:14:03 | Twitter for iPhone | |

| | text | |
|------|---|--|
| 1895 | This is Biden. Biden just tripped... 7/10 http... | |
| 1894 | Ermergerd 12/10 https://t.co/PQni2sjPsm | |

| | expanded_urls | rating_numerator | |
|------|---|------------------|--|
| 1895 | https://twitter.com/dog_rates/status/667405339... | 7 | |
| 1894 | https://twitter.com/dog_rates/status/667435689... | 12 | |

| | rating_denominator | name | favorites | retweet_count | |
|------|--------------------|-------|-----------|---------------|--|
| 1895 | 10 | Biden | NaN | 234 | |
| 1894 | 10 | None | NaN | 89 | |

| | jpg_url | dog_type | breed | |
|------|---|----------|---------------|--|
| 1895 | https://pbs.twimg.com/media/CUMZnmhUEAEBtis.jpg | None | Saint_Bernard | |
| 1894 | https://pbs.twimg.com/media/CUM10HCW4AEgGSi.jpg | None | Rottweiler | |

| | confidence |
|------|------------|
| 1895 | 0.95 |
| 1894 | 0.95 |

```

1895      0.38
1894      1.00

```

```
In [101]: df1.head(2)
```

```

Out[101]:
      tweet_id      timestamp      source \
1895  667405339315146752  2015-11-19 18:13:27  Twitter for iPhone
1894  667435689202614272  2015-11-19 20:14:03  Twitter for iPhone

      text \
1895  This is Biden. Biden just tripped... 7/10 http...
1894      Ermergerd 12/10 https://t.co/PQni2sjPsm

      expanded_urls  rating_numerator \
1895  https://twitter.com/dog_rates/status/667405339...      7
1894  https://twitter.com/dog_rates/status/667435689...      12

      rating_denominator  name  favorites  retweet_count \
1895      10  Biden      NaN      234
1894      10  None      NaN      89

      jpg_url  dog_type      breed \
1895  https://pbs.twimg.com/media/CUMZnmhUEAEbtis.jpg      None  Saint_Bernard
1894  https://pbs.twimg.com/media/CUM10HCW4AEgGSi.jpg      None  Rottweiler

      confidence
1895      0.38
1894      1.00

```

```

In [102]: df1 = df1.rename(columns = {'timestamp': 'tweet_date',
      'source': 'tweet_source',
      'text': 'tweet_text',
      'expanded_urls': 'tweet_url',
      'rating_numerator': 'dog_score',
      'rating_denominator': 'out_of',
      'name': 'dog_name',
      'favorites': 'no_of_likes',
      'jpg_url': 'image_link',
      'breed': 'dog_breed',
      'confidence': 'probability_AI_calc'})

```

```

In [103]: df1 = df1[['tweet_id', 'tweet_date', 'tweet_source', 'tweet_text', 'dog_name',
      'dog_score', 'out_of', 'dog_type', 'dog_breed', 'probability_AI_calc',
      'no_of_likes', 'retweet_count', 'tweet_url', 'image_link']]

```

```
In [104]: df1.head(25)
```

```

Out[104]:
      tweet_id      tweet_date      tweet_source \
1895  667405339315146752  2015-11-19 18:13:27  Twitter for iPhone

```

| | | | | |
|------|--------------------|------------|----------|--------------------|
| 1894 | 667435689202614272 | 2015-11-19 | 20:14:03 | Twitter for iPhone |
| 1893 | 667437278097252352 | 2015-11-19 | 20:20:22 | Twitter for iPhone |
| 1892 | 667443425659232256 | 2015-11-19 | 20:44:47 | Twitter for iPhone |
| 1891 | 667453023279554560 | 2015-11-19 | 21:22:56 | Twitter Web Client |
| 1890 | 667455448082227200 | 2015-11-19 | 21:32:34 | Twitter Web Client |
| 1889 | 667470559035432960 | 2015-11-19 | 22:32:36 | Twitter Web Client |
| 1888 | 667491009379606528 | 2015-11-19 | 23:53:52 | Twitter Web Client |
| 1887 | 667495797102141441 | 2015-11-20 | 00:12:54 | Twitter Web Client |
| 1886 | 667502640335572993 | 2015-11-20 | 00:40:05 | Twitter Web Client |
| 1885 | 667509364010450944 | 2015-11-20 | 01:06:48 | Twitter Web Client |
| 1884 | 667517642048163840 | 2015-11-20 | 01:39:42 | Twitter Web Client |
| 1896 | 667393430834667520 | 2015-11-19 | 17:26:08 | Twitter for iPhone |
| 1883 | 667524857454854144 | 2015-11-20 | 02:08:22 | Twitter Web Client |
| 1881 | 667534815156183040 | 2015-11-20 | 02:47:56 | Twitter Web Client |
| 1880 | 667538891197542400 | 2015-11-20 | 03:04:08 | Twitter Web Client |
| 1879 | 667544320556335104 | 2015-11-20 | 03:25:43 | Twitter Web Client |
| 1878 | 667546741521195010 | 2015-11-20 | 03:35:20 | Twitter Web Client |
| 1877 | 667549055577362432 | 2015-11-20 | 03:44:31 | Twitter Web Client |
| 1876 | 667724302356258817 | 2015-11-20 | 15:20:54 | Twitter Web Client |
| 1875 | 667728196545200128 | 2015-11-20 | 15:36:22 | Twitter Web Client |
| 1874 | 667766675769573376 | 2015-11-20 | 18:09:16 | Twitter Web Client |
| 1873 | 667773195014021121 | 2015-11-20 | 18:35:10 | Twitter Web Client |
| 1872 | 667782464991965184 | 2015-11-20 | 19:12:01 | Twitter for iPhone |
| 1871 | 667793409583771648 | 2015-11-20 | 19:55:30 | Twitter for iPhone |

| | tweet_text | dog_name | dog_score \ |
|------|---|----------|-------------|
| 1895 | This is Biden. Biden just tripped... 7/10 http... | Biden | 7 |
| 1894 | Ermergerd 12/10 https://t.co/PQni2sjPsm | None | 12 |
| 1893 | Never seen this breed before. Very pointy pup... | None | 10 |
| 1892 | Exotic dog here. Long neck. Weird paws. Obsess... | None | 6 |
| 1891 | Meet Cupcake. I would do unspeakable things fo... | Cupcake | 11 |
| 1890 | This is Reese and Twips. Reese protects Twips... | Reese | 7 |
| 1889 | This is a northern Wahoo named Kohl. He runs t... | None | 11 |
| 1888 | Two dogs in this one. Both are rare Jujitsu Py... | None | 7 |
| 1887 | This is Philippe from Soviet Russia. Commandin... | Philippe | 9 |
| 1886 | Say hello to Hall and Oates. Oates is winking ... | Hall | 11 |
| 1885 | This a Norwegian Pewterschmidt named Tickles. ... | None | 12 |
| 1884 | This is Dook & Milo. Dook is struggling to... | Dook | 8 |
| 1896 | This is Fwed. He is a Canadian Asian Taylormad... | Fwed | 8 |
| 1883 | Another topnotch dog. His name is Big Jumpy Ra... | None | 12 |
| 1881 | This is Frank (pronounced "Fronq"). Too many b... | Frank | 8 |
| 1880 | This is a southwest Coriander named Klint. Hat... | None | 9 |
| 1879 | This is Kial. Kial is either wearing a cape, w... | Kial | 10 |
| 1878 | Here is George. George took a selfie of his ne... | George | 9 |
| 1877 | Never seen dog like this. Breathes heavy. Tilt... | None | 1 |
| 1876 | What a dog to start the day with. Very calm. L... | None | 7 |
| 1875 | Meet Olive. He comes to spot by tree to remini... | Olive | 11 |
| 1874 | This is Calvin. He is a Luxembourgian Mayo. Ha... | Calvin | 9 |

| | | | |
|------|---|------|---|
| 1873 | This is a rare Hungarian Pinot named Jessiga. ... | None | 8 |
| 1872 | Super rare dog. Endangered (?). Thinks it's fu... | None | 9 |
| 1871 | Dogs only please. Small cows and other non can... | None | 8 |

| | out_of | dog_type | dog_breed | probability_AI_calc | \ |
|------|--------|----------|-----------------------------|---------------------|---|
| 1895 | 10 | None | Saint_Bernard | 0.38 | |
| 1894 | 10 | None | Rottweiler | 1.00 | |
| 1893 | 10 | None | Unknown | 0.00 | |
| 1892 | 10 | None | Unknown | 0.00 | |
| 1891 | 10 | None | Labrador_retriever | 0.83 | |
| 1890 | 10 | None | Tibetan_terrier | 0.68 | |
| 1889 | 10 | None | toy_poodle | 0.30 | |
| 1888 | 10 | None | borzoi | 0.85 | |
| 1887 | 10 | None | Chihuahua | 0.14 | |
| 1886 | 10 | None | Labrador_retriever | 1.00 | |
| 1885 | 10 | None | beagle | 0.64 | |
| 1884 | 10 | None | Italian_greyhound | 0.13 | |
| 1896 | 10 | None | papillon | 0.56 | |
| 1883 | 10 | None | Chesapeake_Bay_retriever | 0.09 | |
| 1881 | 10 | None | Pembroke | 0.44 | |
| 1880 | 10 | None | Yorkshire_terrier | 0.62 | |
| 1879 | 10 | None | Pomeranian | 0.41 | |
| 1878 | 10 | None | toy_poodle | 0.79 | |
| 1877 | 10 | None | Unknown | 0.00 | |
| 1876 | 10 | None | Unknown | 0.00 | |
| 1875 | 10 | None | kuvasz | 0.36 | |
| 1874 | 10 | None | Unknown | 0.00 | |
| 1873 | 10 | None | West_Highland_white_terrier | 0.36 | |
| 1872 | 10 | None | Unknown | 0.00 | |
| 1871 | 10 | None | dalmatian | 0.54 | |

| | no_of_likes | retweet_count | \ |
|------|-------------|---------------|---|
| 1895 | NaN | 234 | |
| 1894 | NaN | 89 | |
| 1893 | NaN | 257 | |
| 1892 | NaN | 620 | |
| 1891 | NaN | 96 | |
| 1890 | NaN | 66 | |
| 1889 | NaN | 102 | |
| 1888 | NaN | 242 | |
| 1887 | NaN | 294 | |
| 1886 | NaN | 231 | |
| 1885 | NaN | 2272 | |
| 1884 | NaN | 203 | |
| 1896 | NaN | 60 | |
| 1883 | NaN | 1198 | |
| 1881 | NaN | 576 | |
| 1880 | NaN | 72 | |

| | | |
|------|-----|------|
| 1879 | NaN | 568 |
| 1878 | NaN | 138 |
| 1877 | NaN | 2454 |
| 1876 | NaN | 341 |
| 1875 | NaN | 162 |
| 1874 | NaN | 243 |
| 1873 | NaN | 61 |
| 1872 | NaN | 261 |
| 1871 | NaN | 358 |

| | tweet_url \ |
|------|---|
| 1895 | https://twitter.com/dog_rates/status/667405339... |
| 1894 | https://twitter.com/dog_rates/status/667435689... |
| 1893 | https://twitter.com/dog_rates/status/667437278... |
| 1892 | https://twitter.com/dog_rates/status/667443425... |
| 1891 | https://twitter.com/dog_rates/status/667453023... |
| 1890 | https://twitter.com/dog_rates/status/667455448... |
| 1889 | https://twitter.com/dog_rates/status/667470559... |
| 1888 | https://twitter.com/dog_rates/status/667491009... |
| 1887 | https://twitter.com/dog_rates/status/667495797... |
| 1886 | https://twitter.com/dog_rates/status/667502640... |
| 1885 | https://twitter.com/dog_rates/status/667509364... |
| 1884 | https://twitter.com/dog_rates/status/667517642... |
| 1896 | https://twitter.com/dog_rates/status/667393430... |
| 1883 | https://twitter.com/dog_rates/status/667524857... |
| 1881 | https://twitter.com/dog_rates/status/667534815... |
| 1880 | https://twitter.com/dog_rates/status/667538891... |
| 1879 | https://twitter.com/dog_rates/status/667544320... |
| 1878 | https://twitter.com/dog_rates/status/667546741... |
| 1877 | https://twitter.com/dog_rates/status/667549055... |
| 1876 | https://twitter.com/dog_rates/status/667724302... |
| 1875 | https://twitter.com/dog_rates/status/667728196... |
| 1874 | https://twitter.com/dog_rates/status/667766675... |
| 1873 | https://twitter.com/dog_rates/status/667773195... |
| 1872 | https://twitter.com/dog_rates/status/667782464... |
| 1871 | https://twitter.com/dog_rates/status/667793409... |

| | image_link |
|------|---|
| 1895 | https://pbs.twimg.com/media/CUMZnmhUEAEbtis.jpg |
| 1894 | https://pbs.twimg.com/media/CUM10HCW4AEgGSi.jpg |
| 1893 | https://pbs.twimg.com/media/CUM2qWaWoAUZO6L.jpg |
| 1892 | https://pbs.twimg.com/media/CUM8QZwW4AAVsBl.jpg |
| 1891 | https://pbs.twimg.com/media/CUNE_OSUwAAAdHhX.jpg |
| 1890 | https://pbs.twimg.com/media/CUNHMXtU8AAS3HH.jpg |
| 1889 | https://pbs.twimg.com/media/CUNU78YWEAECompB.jpg |
| 1888 | https://pbs.twimg.com/media/CUNniSlUYAEj1Jl.jpg |
| 1887 | https://pbs.twimg.com/media/CUNr4-7UwAAg2lq.jpg |
| 1886 | https://pbs.twimg.com/media/CUNyHTMUyAAQVch.jpg |

```

1885 https://pbs.twimg.com/media/CUN40r5UAAAa5K4.jpg
1884 https://pbs.twimg.com/media/CUN_wiBUkAAakT0.jpg
1896 https://pbs.twimg.com/media/CUM0yd3XIAA113H.jpg
1883 https://pbs.twimg.com/media/CUOGUfJW4AA_eni.jpg
1881 https://pbs.twimg.com/media/CUOPYI5UcAAj_n0.jpg
1880 https://pbs.twimg.com/media/CUOTFZOW4AABsfW.jpg
1879 https://pbs.twimg.com/media/CUOYBbbWIAAXQGU.jpg
1878 https://pbs.twimg.com/media/CUOaOWXWcAAO_Jy.jpg
1877 https://pbs.twimg.com/media/CUOcVCwWsAERUKY.jpg
1876 https://pbs.twimg.com/media/CUQ7tv3W4AA3K1I.jpg
1875 https://pbs.twimg.com/media/CUQ_QahUAAAVQjn.jpg
1874 https://pbs.twimg.com/media/CURiQMnUAAAPT2M.jpg
1873 https://pbs.twimg.com/media/CURoLrOVEAAaWdR.jpg
1872 https://pbs.twimg.com/media/CURwm3cUkAARcO6.jpg
1871 https://pbs.twimg.com/media/CUR6jqVWsAEgGot.jpg

```

```
In [105]: df1.to_csv('twitter_archive_master.csv', encoding='utf-8', index=False)
```

```
In [106]: df_clean = pd.read_csv('twitter_archive_master.csv')
```

```
In [107]: df_clean.head()
```

```

Out[107]:
      tweet_id      tweet_date      tweet_source \
0  667405339315146752  2015-11-19 18:13:27  Twitter for iPhone
1  667435689202614272  2015-11-19 20:14:03  Twitter for iPhone
2  667437278097252352  2015-11-19 20:20:22  Twitter for iPhone
3  667443425659232256  2015-11-19 20:44:47  Twitter for iPhone
4  667453023279554560  2015-11-19 21:22:56  Twitter Web Client

      tweet_text      dog_name      dog_score \
0  This is Biden. Biden just tripped... 7/10 http...  Biden          7
1      Ermergerd 12/10 https://t.co/PQni2sjPsm      None          12
2  Never seen this breed before. Very pointy pup...      None          10
3  Exotic dog here. Long neck. Weird paws. Obsess...      None           6
4  Meet Cupcake. I would do unspeakable things fo...  Cupcake          11

      out_of      dog_type      dog_breed      probability_AI_calc      no_of_likes \
0      10      None      Saint_Bernard          0.38          NaN
1      10      None      Rottweiler          1.00          NaN
2      10      None      Unknown          0.00          NaN
3      10      None      Unknown          0.00          NaN
4      10      None  Labrador_retriever          0.83          NaN

      retweet_count      tweet_url \
0      234  https://twitter.com/dog_rates/status/667405339...
1      89  https://twitter.com/dog_rates/status/667435689...
2      257  https://twitter.com/dog_rates/status/667437278...
3      620  https://twitter.com/dog_rates/status/667443425...
4      96  https://twitter.com/dog_rates/status/667453023...

```

```

                                image_link
0  https://pbs.twimg.com/media/CUMZnmhUEAEbtis.jpg
1  https://pbs.twimg.com/media/CUM10HCW4AEgGSi.jpg
2  https://pbs.twimg.com/media/CUM2qWaWoAUZ06L.jpg
3  https://pbs.twimg.com/media/CUM8QZwW4AAVsBl.jpg
4  https://pbs.twimg.com/media/CUNE_OSUwAAAdHhX.jpg

```

```

In [108]: import matplotlib.pyplot as plt
          import seaborn as sns

```

```

In [110]: sns.set(style="darkgrid",font_scale=1.5)

```

```

In [111]: %matplotlib inline

```

```

In [112]: df_clean.describe()

```

```

Out[112]:
          tweet_id  dog_score  out_of  probability_AI_calc \
count  1.971000e+03  1971.000000  1971.000000  1971.000000
mean    7.360418e+17   12.207509   10.457128    0.466301
std     6.752810e+16   41.623366    6.788656    0.339697
min     6.660209e+17    0.000000   10.000000    0.000000
25%     6.758656e+17   10.000000   10.000000    0.140000
50%     7.088343e+17   11.000000   10.000000    0.460000
75%     7.880951e+17   12.000000   10.000000    0.780000
max     8.924206e+17  1776.000000  170.000000    1.000000

```

```

          no_of_likes  retweet_count
count              0.0    1971.000000
mean              NaN    2784.449518
std              NaN    4697.662893
min              NaN     16.000000
25%              NaN     628.500000
50%              NaN    1367.000000
75%              NaN    3239.000000
max              NaN   79515.000000

```

```

In [113]: df_clean.dog_breed.value_counts().nlargest(10)

```

```

Out[113]:
Unknown          305
golden_retriever  156
Labrador_retriever 106
Pembroke         94
Chihuahua        90
pug              62
toy_poodle       50
chow            48
Samoyed          42
Pomeranian       41
Name: dog_breed, dtype: int64

```

```
In [114]: df_clean.loc[df_clean['dog_breed'] == 'Unknown', 'dog_breed'] = None
```

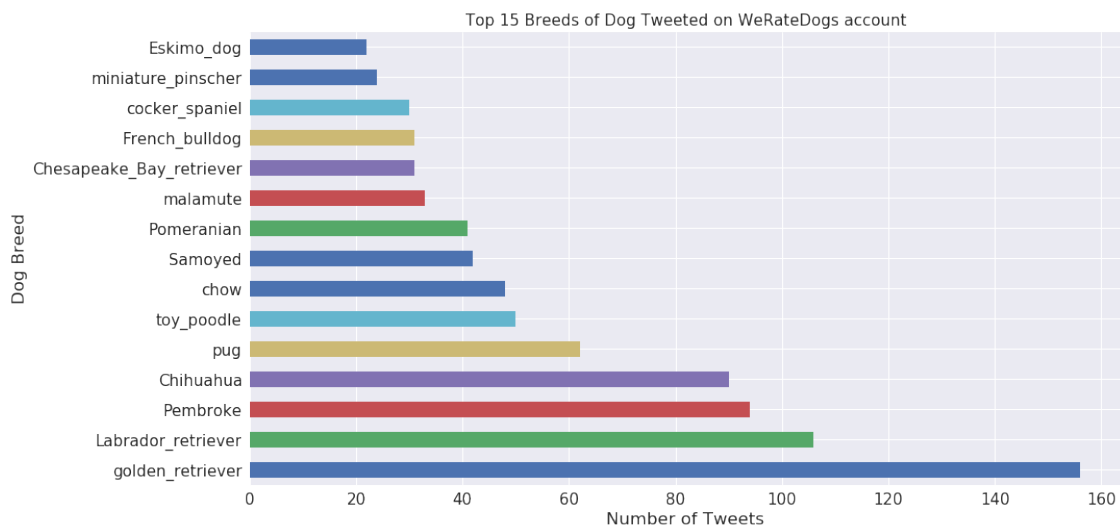
```
In [115]: df_clean.dog_breed.count()
```

```
Out[115]: 1666
```

```
In [116]: df_clean.dog_breed.value_counts().nlargest(10)
```

```
Out[116]: golden_retriever      156
Labrador_retriever      106
Pembroke                 94
Chihuahua                90
pug                     62
toy_poodle               50
chow                    48
Samoyed                 42
Pomeranian              41
malamute                33
Name: dog_breed, dtype: int64
```

```
In [117]: fig = plt.figure(figsize=(15,8))
df_clean.groupby('dog_breed').count()['dog_score'].sort_values(ascending=False).nlarge
plt.title("Top 15 Breeds of Dog Tweeted on WeRateDogs account",fontsize=15)
plt.ylabel("Dog Breed")
plt.xlabel("Number of Tweets");
```



```
In [120]: df_filtered = df_clean[df_clean['dog_score'] <= 14]
```

```
In [121]: df_filtered['dog_score'].describe()
```

```

Out[121]: count    1955.000000
          mean      10.538619
          std       2.175258
          min       0.000000
          25%      10.000000
          50%      11.000000
          75%      12.000000
          max      14.000000
          Name: dog_score, dtype: float64

```

```

In [122]: df_filtered.groupby('dog_breed')['dog_score'].describe()

```

```

Out[122]:
          count      mean      std  min  25%  50%  \
dog_breed
Afghan_hound      3.0   9.666667  5.773503   3.0   8.00  13.0
Airedale          12.0   9.833333  2.249579   6.0   9.25  10.5
American_Staffordshire_terrier  16.0  10.312500  3.070695   0.0  10.00  11.0
Appenzeller        2.0  11.000000  2.828427   9.0  10.00  11.0
Australian_terrier  2.0  11.500000  2.121320  10.0  10.75  11.5
Bedlington_terrier  6.0  10.333333  3.141125   6.0   8.25  10.5
Bernese_mountain_dog  11.0  11.272727  1.103713  10.0  10.00  12.0
Blenheim_spaniel   10.0  11.100000  1.100505   9.0  10.25  11.5
Border_collie      12.0  10.583333  2.503028   4.0   9.75  11.5
Border_terrier      7.0  12.142857  0.690066  11.0  12.00  12.0
Boston_bull        12.0   9.416667  2.108784   5.0   8.75  10.0
Bouvier_des_Flandres    1.0  13.000000      NaN  13.0  13.00  13.0
Brabancon_griffon     3.0  10.000000  0.000000  10.0  10.00  10.0
Brittany_spaniel     8.0  10.375000  1.685018   7.0   9.75  11.0
Cardigan           21.0  11.142857  1.590148   7.0  10.00  11.0
Chesapeake_Bay_retriever  31.0  10.838710  1.439833   8.0  10.00  10.0
Chihuahua          90.0  10.500000  2.089299   3.0  10.00  11.0
Dandie_Dinmont      7.0   9.571429  3.258688   3.0   9.00  11.0
Doberman           9.0  11.333333  1.000000  10.0  11.00  11.0
English_setter      8.0   9.875000  1.726888   8.0   8.75   9.5
English_springer    10.0  10.500000  2.223611   5.0  10.00  10.5
EntleBucher        1.0  11.000000      NaN  11.0  11.00  11.0
Eskimo_dog         22.0  11.409091  1.402688   9.0  10.00  12.0
French_bulldog     31.0  11.193548  1.796652   8.0  10.00  12.0
German_shepherd     21.0  11.000000  1.449138   8.0  10.00  11.0
German_short-haired_pointer    8.0  10.500000  0.755929  10.0  10.00  10.0
Gordon_setter       4.0  11.750000  2.217356   9.0  10.50  12.0
Great_Dane         11.0  10.818182  1.940009   7.0  10.00  11.0
Great_Pyrenees     14.0  11.428571  1.222500   9.0  11.00  11.5
Greater_Swiss_Mountain_dog    3.0  11.333333  2.081666   9.0  10.50  12.0
Ibizan_hound       4.0   9.250000  0.957427   8.0   8.75   9.5
Irish_setter        4.0  11.750000  1.707825  10.0  10.75  11.5
Irish_terrier       6.0  10.000000  2.280351   6.0  10.00  10.0
Irish_water_spaniel    3.0  11.333333  1.154701  10.0  11.00  12.0

```

| | | | | | | |
|--------------------------------|-------|-----------|----------|------|-------|------|
| Irish_wolfhound | 1.0 | 11.000000 | NaN | 11.0 | 11.00 | 11.0 |
| Italian_greyhound | 17.0 | 10.000000 | 1.870829 | 6.0 | 8.00 | 11.0 |
| Japanese_spaniel | 1.0 | 5.000000 | NaN | 5.0 | 5.00 | 5.0 |
| ... | ... | ... | ... | ... | ... | ... |
| bluetick | 4.0 | 10.500000 | 1.914854 | 8.0 | 9.50 | 11.0 |
| borzoi | 9.0 | 10.111111 | 2.204793 | 7.0 | 9.00 | 10.0 |
| boxer | 11.0 | 10.545455 | 1.572491 | 8.0 | 10.00 | 10.0 |
| briard | 3.0 | 12.333333 | 0.577350 | 12.0 | 12.00 | 12.0 |
| bull_mastiff | 5.0 | 10.400000 | 1.140175 | 9.0 | 10.00 | 10.0 |
| cairn | 3.0 | 11.000000 | 1.000000 | 10.0 | 10.50 | 11.0 |
| chow | 47.0 | 11.404255 | 1.361897 | 7.0 | 11.00 | 12.0 |
| cocker_spaniel | 30.0 | 11.333333 | 1.184187 | 9.0 | 10.25 | 11.5 |
| collie | 11.0 | 10.909091 | 1.221028 | 9.0 | 10.00 | 11.0 |
| curly-coated_retriever | 3.0 | 10.333333 | 2.309401 | 9.0 | 9.00 | 9.0 |
| dalmatian | 12.0 | 9.333333 | 2.188122 | 5.0 | 8.00 | 9.5 |
| flat-coated_retriever | 8.0 | 10.875000 | 2.531939 | 5.0 | 10.75 | 12.0 |
| giant_schnauzer | 3.0 | 11.000000 | 1.732051 | 10.0 | 10.00 | 10.0 |
| golden_retriever | 155.0 | 11.561290 | 1.324513 | 5.0 | 11.00 | 12.0 |
| groenendael | 2.0 | 10.000000 | 0.000000 | 10.0 | 10.00 | 10.0 |
| keeshond | 4.0 | 10.250000 | 1.500000 | 9.0 | 9.00 | 10.0 |
| kelpie | 13.0 | 11.307692 | 1.109400 | 10.0 | 11.00 | 11.0 |
| komondor | 3.0 | 11.000000 | 1.000000 | 10.0 | 10.50 | 11.0 |
| kuvasz | 16.0 | 10.625000 | 2.276694 | 4.0 | 10.00 | 11.5 |
| malamute | 33.0 | 10.878788 | 1.218544 | 8.0 | 10.00 | 11.0 |
| malinois | 11.0 | 10.818182 | 1.990888 | 6.0 | 10.00 | 11.0 |
| miniature_pinscher | 24.0 | 10.041667 | 2.612331 | 2.0 | 9.75 | 11.0 |
| miniature_poodle | 8.0 | 9.875000 | 1.552648 | 8.0 | 8.75 | 10.0 |
| miniature_schnauzer | 5.0 | 9.600000 | 2.792848 | 5.0 | 9.00 | 11.0 |
| papillon | 7.0 | 10.000000 | 1.632993 | 8.0 | 8.50 | 11.0 |
| pug | 62.0 | 10.241935 | 1.816910 | 3.0 | 9.25 | 10.0 |
| redbone | 6.0 | 9.666667 | 2.250926 | 7.0 | 7.75 | 10.0 |
| schipperke | 13.0 | 10.923077 | 1.382120 | 8.0 | 10.00 | 11.0 |
| silky_terrier | 1.0 | 12.000000 | NaN | 12.0 | 12.00 | 12.0 |
| soft-coated_wheaten_terrier | 12.0 | 8.166667 | 3.128559 | 3.0 | 5.50 | 9.5 |
| standard_poodle | 11.0 | 10.727273 | 1.678744 | 9.0 | 9.50 | 10.0 |
| standard_schnauzer | 1.0 | 12.000000 | NaN | 12.0 | 12.00 | 12.0 |
| toy_poodle | 50.0 | 11.000000 | 1.245400 | 7.0 | 10.00 | 11.0 |
| toy_terrier | 3.0 | 10.333333 | 1.527525 | 9.0 | 9.50 | 10.0 |
| vizsla | 13.0 | 10.846154 | 1.281025 | 8.0 | 11.00 | 11.0 |
| whippet | 11.0 | 10.181818 | 1.834022 | 8.0 | 8.50 | 10.0 |
| wire-haired_fox_terrier | 2.0 | 11.500000 | 0.707107 | 11.0 | 11.25 | 11.5 |
| | | | | | | |
| | 75% | max | | | | |
| dog_breed | | | | | | |
| Afghan_hound | 13.00 | 13.0 | | | | |
| Airedale | 11.25 | 12.0 | | | | |
| American_Staffordshire_terrier | 12.00 | 13.0 | | | | |
| Appenzeller | 12.00 | 13.0 | | | | |

| | | |
|-----------------------------|-------|------|
| Australian_terrier | 12.25 | 13.0 |
| Bedlington_terrier | 12.75 | 14.0 |
| Bernese_mountain_dog | 12.00 | 13.0 |
| Blenheim_spaniel | 12.00 | 12.0 |
| Border_collie | 12.00 | 13.0 |
| Border_terrier | 12.50 | 13.0 |
| Boston_bull | 11.00 | 12.0 |
| Bouvier_des_Flandres | 13.00 | 13.0 |
| Brabancon_griffon | 10.00 | 10.0 |
| Brittany_spaniel | 11.25 | 12.0 |
| Cardigan | 12.00 | 13.0 |
| Chesapeake_Bay_retriever | 12.00 | 13.0 |
| Chihuahua | 12.00 | 14.0 |
| Dandie_Dinmont | 11.00 | 13.0 |
| Doberman | 12.00 | 13.0 |
| English_setter | 11.00 | 13.0 |
| English_springer | 12.00 | 13.0 |
| EntleBucher | 11.00 | 11.0 |
| Eskimo_dog | 12.00 | 14.0 |
| French_bulldog | 12.00 | 14.0 |
| German_shepherd | 12.00 | 13.0 |
| German_short-haired_pointer | 11.00 | 12.0 |
| Gordon_setter | 13.25 | 14.0 |
| Great_Dane | 12.00 | 13.0 |
| Great_Pyrenees | 12.00 | 13.0 |
| Greater_Swiss_Mountain_dog | 12.50 | 13.0 |
| Ibizan_hound | 10.00 | 10.0 |
| Irish_setter | 12.50 | 14.0 |
| Irish_terrier | 10.75 | 13.0 |
| Irish_water_spaniel | 12.00 | 12.0 |
| Irish_wolfhound | 11.00 | 11.0 |
| Italian_greyhound | 11.00 | 13.0 |
| Japanese_spaniel | 5.00 | 5.0 |
| ... | ... | ... |
| bluetick | 12.00 | 12.0 |
| borzoi | 11.00 | 13.0 |
| boxer | 11.50 | 13.0 |
| briard | 12.50 | 13.0 |
| bull_mastiff | 11.00 | 12.0 |
| cairn | 11.50 | 12.0 |
| chow | 12.00 | 13.0 |
| cocker_spaniel | 12.00 | 13.0 |
| collie | 12.00 | 13.0 |
| curly-coated_retriever | 11.00 | 13.0 |
| dalmatian | 10.25 | 13.0 |
| flat-coated_retriever | 12.00 | 13.0 |
| giant_schnauzer | 11.50 | 13.0 |
| golden_retriever | 12.00 | 14.0 |

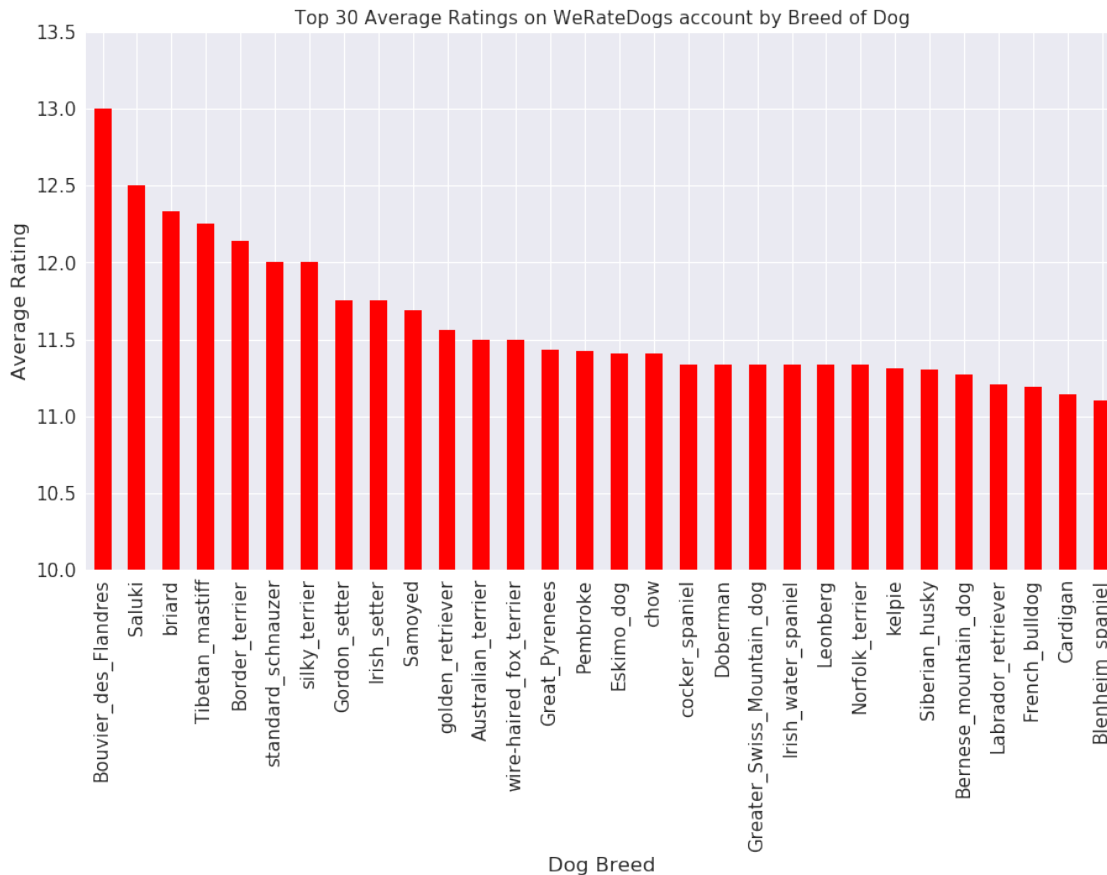
| | | |
|-----------------------------|-------|------|
| groenendael | 10.00 | 10.0 |
| keeshond | 11.25 | 12.0 |
| kelpie | 12.00 | 13.0 |
| komondor | 11.50 | 12.0 |
| kuvasz | 12.00 | 13.0 |
| malamute | 12.00 | 13.0 |
| malinois | 12.00 | 13.0 |
| miniature_pinscher | 12.00 | 12.0 |
| miniature_poodle | 10.50 | 12.0 |
| miniature_schnauzer | 11.00 | 12.0 |
| papillon | 11.00 | 12.0 |
| pug | 11.00 | 13.0 |
| redbone | 11.50 | 12.0 |
| schipperke | 12.00 | 13.0 |
| silky_terrier | 12.00 | 12.0 |
| soft-coated_wheaten_terrier | 10.25 | 12.0 |
| standard_poodle | 11.50 | 14.0 |
| standard_schnauzer | 12.00 | 12.0 |
| toy_poodle | 12.00 | 13.0 |
| toy_terrier | 11.00 | 12.0 |
| vizsla | 11.00 | 13.0 |
| whippet | 12.00 | 13.0 |
| wire-haired_fox_terrier | 11.75 | 12.0 |

[112 rows x 8 columns]

```
In [123]: df_filtered.groupby('dog_breed')['dog_score'].mean().sort_values(ascending=False).nlargest(30)
```

```
Out[123]: dog_breed
Bouvier_des_Flandres    13.000000
Saluki                  12.500000
briard                  12.333333
Tibetan_mastiff         12.250000
Border_terrier          12.142857
standard_schnauzer      12.000000
silky_terrier           12.000000
Gordon_setter           11.750000
Irish_setter            11.750000
Samoyed                 11.690476
Name: dog_score, dtype: float64
```

```
In [125]: fig = plt.figure(figsize=(15,8))
df_filtered.groupby('dog_breed')['dog_score'].mean().sort_values(ascending=False).nlargest(30)
plt.title("Top 30 Average Ratings on WeRateDogs account by Breed of Dog",fontsize=15)
plt.ylabel("Average Rating")
plt.xlabel("Dog Breed")
plt.ylim(10,13.5);
```

```
In [126]: morethan9counts = df_clean['dog_breed'].value_counts()
```

```
In [129]: df_breedsmorethan10ratings = df_clean[df_clean['dog_breed'].isin(morethan9counts[moret
```

```
In [130]: df_breedsmorethan10ratings
```

```
Out[130]:
```

| | tweet_id | tweet_date | tweet_source |
|----|--------------------|---------------------|--------------------|
| 1 | 667435689202614272 | 2015-11-19 20:14:03 | Twitter for iPhone |
| 4 | 667453023279554560 | 2015-11-19 21:22:56 | Twitter Web Client |
| 6 | 667470559035432960 | 2015-11-19 22:32:36 | Twitter Web Client |
| 7 | 667491009379606528 | 2015-11-19 23:53:52 | Twitter Web Client |
| 8 | 667495797102141441 | 2015-11-20 00:12:54 | Twitter Web Client |
| 9 | 667502640335572993 | 2015-11-20 00:40:05 | Twitter Web Client |
| 10 | 667509364010450944 | 2015-11-20 01:06:48 | Twitter Web Client |
| 11 | 667517642048163840 | 2015-11-20 01:39:42 | Twitter Web Client |
| 13 | 667524857454854144 | 2015-11-20 02:08:22 | Twitter Web Client |
| 14 | 667534815156183040 | 2015-11-20 02:47:56 | Twitter Web Client |
| 15 | 667538891197542400 | 2015-11-20 03:04:08 | Twitter Web Client |
| 16 | 667544320556335104 | 2015-11-20 03:25:43 | Twitter Web Client |
| 17 | 667546741521195010 | 2015-11-20 03:35:20 | Twitter Web Client |

| | | | |
|------|--------------------|---------------------|--------------------|
| 20 | 667728196545200128 | 2015-11-20 15:36:22 | Twitter Web Client |
| 22 | 667773195014021121 | 2015-11-20 18:35:10 | Twitter Web Client |
| 24 | 667793409583771648 | 2015-11-20 19:55:30 | Twitter for iPhone |
| 26 | 667530908589760512 | 2015-11-20 02:32:25 | Twitter Web Client |
| 28 | 667211855547486208 | 2015-11-19 05:24:37 | Twitter for iPhone |
| 29 | 667200525029539841 | 2015-11-19 04:39:35 | Twitter for iPhone |
| 30 | 666781792255496192 | 2015-11-18 00:55:42 | Twitter for iPhone |
| 34 | 666826780179869698 | 2015-11-18 03:54:28 | Twitter for iPhone |
| 35 | 666835007768551424 | 2015-11-18 04:27:09 | Twitter for iPhone |
| 38 | 666996132027977728 | 2015-11-18 15:07:24 | Twitter for iPhone |
| 40 | 667044094246576128 | 2015-11-18 18:17:59 | Twitter for iPhone |
| 41 | 667062181243039745 | 2015-11-18 19:29:52 | Twitter for iPhone |
| 43 | 667073648344346624 | 2015-11-18 20:15:26 | Twitter for iPhone |
| 44 | 667090893657276420 | 2015-11-18 21:23:57 | Twitter for iPhone |
| 45 | 667119796878725120 | 2015-11-18 23:18:48 | Twitter for iPhone |
| 46 | 667138269671505920 | 2015-11-19 00:32:12 | Twitter for iPhone |
| 47 | 667152164079423490 | 2015-11-19 01:27:25 | Twitter for iPhone |
| 48 | 667160273090932737 | 2015-11-19 01:59:39 | Twitter for iPhone |
| 49 | 667165590075940865 | 2015-11-19 02:20:46 | Twitter for iPhone |
| 51 | 667174963120574464 | 2015-11-19 02:58:01 | Twitter for iPhone |
| 52 | 667176164155375616 | 2015-11-19 03:02:47 | Twitter for iPhone |
| 53 | 667177989038297088 | 2015-11-19 03:10:02 | Twitter for iPhone |
| 54 | 667182792070062081 | 2015-11-19 03:29:07 | Twitter for iPhone |
| 56 | 667192066997374976 | 2015-11-19 04:05:59 | Twitter for iPhone |
| ... | ... | ... | ... |
| 1925 | 674774481756377088 | 2015-12-10 02:15:47 | Twitter for iPhone |
| 1926 | 675006312288268288 | 2015-12-10 17:37:00 | Twitter for iPhone |
| 1928 | 692752401762250755 | 2016-01-28 16:53:37 | Twitter for iPhone |
| 1929 | 837820167694528512 | 2017-03-04 00:21:08 | Twitter for iPhone |
| 1934 | 693629975228977152 | 2016-01-31 03:00:47 | Twitter for iPhone |
| 1935 | 873213775632977920 | 2017-06-09 16:22:42 | Twitter for iPhone |
| 1938 | 693262851218264065 | 2016-01-30 02:41:58 | Twitter for iPhone |
| 1939 | 692828166163931137 | 2016-01-28 21:54:41 | Twitter for iPhone |
| 1940 | 700151421916807169 | 2016-02-18 02:54:41 | Twitter for iPhone |
| 1941 | 747963614829678593 | 2016-06-29 01:23:16 | Twitter for iPhone |
| 1942 | 673956914389192708 | 2015-12-07 20:07:04 | Twitter for iPhone |
| 1943 | 759793422261743616 | 2016-07-31 16:50:42 | Twitter for iPhone |
| 1944 | 693095443459342336 | 2016-01-29 15:36:45 | Twitter for iPhone |
| 1945 | 674038233588723717 | 2015-12-08 01:30:12 | Twitter for iPhone |
| 1946 | 874296783580663808 | 2017-06-12 16:06:11 | Twitter for iPhone |
| 1947 | 673919437611909120 | 2015-12-07 17:38:09 | Twitter for iPhone |
| 1948 | 878776093423087618 | 2017-06-25 00:45:22 | Twitter for iPhone |
| 1949 | 874012996292530176 | 2017-06-11 21:18:31 | Twitter for iPhone |
| 1950 | 793195938047070209 | 2016-10-31 21:00:23 | Twitter for iPhone |
| 1951 | 802239329049477120 | 2016-11-25 19:55:35 | Twitter for iPhone |
| 1952 | 752519690950500352 | 2016-07-11 15:07:30 | Twitter for iPhone |
| 1953 | 743253157753532416 | 2016-06-16 01:25:36 | Twitter for iPhone |
| 1954 | 803773340896923648 | 2016-11-30 01:31:12 | Twitter for iPhone |

| | | | |
|------|--------------------|---------------------|--------------------|
| 1955 | 825535076884762624 | 2017-01-29 02:44:34 | Twitter for iPhone |
| 1956 | 825026590719483904 | 2017-01-27 17:04:02 | Twitter for iPhone |
| 1957 | 744995568523612160 | 2016-06-20 20:49:19 | Twitter for iPhone |
| 1958 | 822872901745569793 | 2017-01-21 18:26:02 | Twitter for iPhone |
| 1959 | 889665388333682689 | 2017-07-25 01:55:32 | Twitter for iPhone |
| 1960 | 889531135344209921 | 2017-07-24 17:02:04 | Twitter for iPhone |
| 1961 | 867421006826221569 | 2017-05-24 16:44:18 | Twitter for iPhone |
| 1962 | 819952236453363712 | 2017-01-13 17:00:21 | Twitter for iPhone |
| 1963 | 738537504001953792 | 2016-06-03 01:07:16 | Twitter for iPhone |
| 1965 | 787717603741622272 | 2016-10-16 18:11:26 | Twitter for iPhone |
| 1966 | 859607811541651456 | 2017-05-03 03:17:27 | Twitter for iPhone |
| 1967 | 790946055508652032 | 2016-10-25 16:00:09 | Twitter for iPhone |
| 1968 | 756275833623502848 | 2016-07-21 23:53:04 | Twitter for iPhone |
| 1970 | 751132876104687617 | 2016-07-07 19:16:47 | Twitter for iPhone |

| | tweet_text | dog_name | dog_score \ |
|----|--|-----------|-------------|
| 1 | Ermergerd 12/10 https://t.co/PQni2sjPsm | None | 12 |
| 4 | Meet Cupcake. I would do unspeakable things fo... | Cupcake | 11 |
| 6 | This is a northern Wahoo named Kohl. He runs t... | None | 11 |
| 7 | Two dogs in this one. Both are rare Jujitsu Py... | None | 7 |
| 8 | This is Philippe from Soviet Russia. Commandin... | Philippe | 9 |
| 9 | Say hello to Hall and Oates. Oates is winking ... | Hall | 11 |
| 10 | This a Norwegian Pewterschmidt named Tickle... | None | 12 |
| 11 | This is Dook & Milo. Dook is struggling to... | Dook | 8 |
| 13 | Another topnotch dog. His name is Big Jumpy Ra... | None | 12 |
| 14 | This is Frank (pronounced "Fronq"). Too many b... | Frank | 8 |
| 15 | This is a southwest Coriander named Klint. Hat... | None | 9 |
| 16 | This is Kial. Kial is either wearing a cape, w... | Kial | 10 |
| 17 | Here is George. George took a selfie of his ne... | George | 9 |
| 20 | Meet Olive. He comes to spot by tree to remini... | Olive | 11 |
| 22 | This is a rare Hungarian Pinot named Jessiga. ... | None | 8 |
| 24 | Dogs only please. Small cows and other non can... | None | 8 |
| 26 | Meet Naphaniel. He doesn't necessarily enjoy h... | Naphaniel | 10 |
| 28 | This is Genevieve. She is a golden retriever c... | Genevieve | 9 |
| 29 | This is Joshwa. He is a fuckboy supreme. He cl... | Joshwa | 11 |
| 30 | This is a purebred Bacardi named Octaviath. Ca... | None | 10 |
| 34 | 12/10 simply brilliant pup https://t.co/V6ZzG4... | None | 12 |
| 35 | These are Peruvian Feldspars. Their names are ... | None | 10 |
| 38 | This is Carll. He wants to be a donkey. But al... | Carll | 10 |
| 40 | 12/10 gimme now https://t.co/QZAnwgnOMB | None | 12 |
| 41 | This is Keet. He is a Floridian Amukamara. Abs... | Keet | 10 |
| 43 | Here is Dave. He is actually just a skinny leg... | Dave | 10 |
| 44 | This is Clybe. He is an Anemone Valdez. One ea... | Clybe | 7 |
| 45 | This is Gabe. He is a southern Baklava. Gabe h... | Gabe | 10 |
| 46 | Extremely intelligent dog here. Has learned to... | None | 10 |
| 47 | This is Pipsy. He is a fluffball. Enjoys trave... | Pipsy | 12 |
| 48 | This is Bradlay. He is a Ronaldinho Matsuyama ... | Bradlay | 11 |
| 49 | This is Churlie. AKA Fetty Woof. Lost eye savi... | Churlie | 10 |

| | | | |
|------|--|-----------|-----|
| 51 | This is Clarence. His face says he doesn't wan... | Clarence | 9 |
| 52 | These are strange dogs. All have toupees. Long... | None | 4 |
| 53 | This is a Dasani Kingfisher from Maine. His na... | None | 8 |
| 54 | This is Timison. He just told an awful joke bu... | Timison | 10 |
| 56 | *takes several long deep breaths* omg omg oMG ... | None | 12 |
| ... | ... | ... | ... |
| 1925 | This is Superpup. His head isn't proportional ... | Superpup | 11 |
| 1926 | Say hello to Mollie. This pic was taken after ... | Mollie | 10 |
| 1928 | "Hello yes could I get one pupper to go please... | None | 13 |
| 1929 | Here's a pupper before and after being asked "... | None | 12 |
| 1934 | This pupper is afraid of its own feet. 12/10 w... | None | 12 |
| 1935 | This is Sierra. She's one precious pupper. Abs... | Sierra | 12 |
| 1938 | I hope you guys enjoy this beautiful snowy pup... | None | 11 |
| 1939 | This pupper just descended from heaven. 12/10 ... | None | 12 |
| 1940 | If a pupper gave that to me I'd probably start... | None | 11 |
| 1941 | PUPPER NOOOOOO BEHIND YOUUU 10/10 pls keep this... | None | 10 |
| 1942 | This is one esteemed pupper. Just graduated co... | None | 10 |
| 1943 | Meet Maggie & Lila. Maggie is the doggo, L... | Maggie | 12 |
| 1944 | This is Milo. He doesn't understand your fancy... | Milo | 10 |
| 1945 | This is Kaiya. She's an aspiring shoe model. 1... | Kaiya | 12 |
| 1946 | This is Jed. He may be the fanciest pupper in ... | Jed | 13 |
| 1947 | This is Obie. He is on guard watching for evil... | Obie | 11 |
| 1948 | This is Snoopy. He's a proud #PrideMonthPuppo... | Snoopy | 13 |
| 1949 | This is Sebastian. He can't see all the colors... | Sebastian | 13 |
| 1950 | Say hello to Lily. She's pupset that her costu... | Lily | 12 |
| 1951 | This is Loki. He'll do your taxes for you. Can... | Loki | 12 |
| 1952 | Hopefully this puppo on a swing will help get ... | None | 11 |
| 1953 | This is Kilo. He cannot reach the snackum. Nif... | Kilo | 10 |
| 1954 | This is Diogi. He fell in the pool as soon as ... | Diogi | 12 |
| 1955 | Here's a very loving and accepting puppo. Appe... | None | 14 |
| 1956 | Say hello to Pablo. He's one gorgeous puppo. A... | Pablo | 12 |
| 1957 | This is Abby. She got her face stuck in a glas... | Abby | 9 |
| 1958 | Here's a super supportive puppo participating ... | None | 13 |
| 1959 | Here's a puppo that seems to be on the fence a... | None | 13 |
| 1960 | This is Stuart. He's sporting his favorite fan... | Stuart | 13 |
| 1961 | This is Shikha. She just watched you drop a sk... | Shikha | 12 |
| 1962 | This is Oliver. He has dreams of being a servi... | Oliver | 13 |
| 1963 | This is Bayley. She fell asleep trying to esca... | Bayley | 11 |
| 1965 | This is Tonks. She is a service puppo. Can hea... | Tonks | 13 |
| 1966 | Sorry for the lack of posts today. I came home... | None | 13 |
| 1967 | This is Betty. She's assisting with the dishes... | Betty | 12 |
| 1968 | When ur older siblings get to play in the deep... | None | 10 |
| 1970 | This is Cooper. He's just so damn happy. 10/10... | Cooper | 10 |

| | out_of | dog_type | dog_breed | probability_AI_calc \ |
|---|--------|----------|--------------------|-----------------------|
| 1 | 10 | None | Rottweiler | 1.00 |
| 4 | 10 | None | Labrador_retriever | 0.83 |
| 6 | 10 | None | toy_poodle | 0.30 |

| | | | | |
|------|-----|--------|-----------------------------|------|
| 7 | 10 | None | borzoi | 0.85 |
| 8 | 10 | None | Chihuahua | 0.14 |
| 9 | 10 | None | Labrador_retriever | 1.00 |
| 10 | 10 | None | beagle | 0.64 |
| 11 | 10 | None | Italian_greyhound | 0.13 |
| 13 | 10 | None | Chesapeake_Bay_retriever | 0.09 |
| 14 | 10 | None | Pembroke | 0.44 |
| 15 | 10 | None | Yorkshire_terrier | 0.62 |
| 16 | 10 | None | Pomeranian | 0.41 |
| 17 | 10 | None | toy_poodle | 0.79 |
| 20 | 10 | None | kuvasz | 0.36 |
| 22 | 10 | None | West_Highland_white_terrier | 0.36 |
| 24 | 10 | None | dalmatian | 0.54 |
| 26 | 10 | None | golden_retriever | 0.63 |
| 28 | 10 | None | golden_retriever | 0.46 |
| 29 | 10 | None | Siberian_husky | 0.69 |
| 30 | 10 | None | Italian_greyhound | 0.62 |
| 34 | 10 | None | Maltese_dog | 0.36 |
| 35 | 10 | None | Airedale | 0.45 |
| 38 | 10 | None | Rottweiler | 0.06 |
| 40 | 10 | None | golden_retriever | 0.77 |
| 41 | 10 | None | Chesapeake_Bay_retriever | 0.83 |
| 43 | 10 | None | Chihuahua | 0.48 |
| 44 | 10 | None | Chihuahua | 0.96 |
| 45 | 10 | None | Pembroke | 0.74 |
| 46 | 10 | None | West_Highland_white_terrier | 0.75 |
| 47 | 10 | None | toy_poodle | 0.54 |
| 48 | 10 | None | golden_retriever | 0.47 |
| 49 | 10 | None | miniature_pinscher | 0.14 |
| 51 | 10 | None | toy_poodle | 0.27 |
| 52 | 10 | None | soft-coated_wheaten_terrier | 0.32 |
| 53 | 10 | None | vizsla | 0.26 |
| 54 | 10 | None | golden_retriever | 0.95 |
| 56 | 10 | None | Rottweiler | 0.28 |
| ... | ... | ... | ... | ... |
| 1925 | 10 | pupper | Chihuahua | 0.41 |
| 1926 | 10 | pupper | boxer | 0.65 |
| 1928 | 10 | pupper | Samoyed | 0.47 |
| 1929 | 10 | pupper | golden_retriever | 0.89 |
| 1934 | 10 | pupper | pug | 0.84 |
| 1935 | 10 | pupper | vizsla | 0.62 |
| 1938 | 10 | pupper | golden_retriever | 0.99 |
| 1939 | 10 | pupper | Samoyed | 0.99 |
| 1940 | 10 | pupper | Italian_greyhound | 0.18 |
| 1941 | 10 | pupper | kelpie | 0.31 |
| 1942 | 10 | pupper | pug | 0.59 |
| 1943 | 10 | None | golden_retriever | 0.99 |
| 1944 | 10 | pupper | Yorkshire_terrier | 0.03 |

| | | | | |
|------|----|--------|--------------------------------|------|
| 1945 | 10 | pupper | Eskimo_dog | 0.36 |
| 1946 | 10 | pupper | cocker_spaniel | 0.44 |
| 1947 | 10 | pupper | schipperke | 0.12 |
| 1948 | 10 | puppo | Italian_greyhound | 0.73 |
| 1949 | 10 | puppo | Cardigan | 0.81 |
| 1950 | 10 | puppo | Labrador_retriever | 0.65 |
| 1951 | 10 | puppo | Eskimo_dog | 0.48 |
| 1952 | 10 | puppo | Labrador_retriever | 0.00 |
| 1953 | 10 | puppo | malamute | 0.44 |
| 1954 | 10 | puppo | miniature_pinscher | 0.82 |
| 1955 | 10 | puppo | Rottweiler | 0.68 |
| 1956 | 10 | puppo | Eskimo_dog | 0.52 |
| 1957 | 10 | puppo | Old_English_sheepdog | 0.43 |
| 1958 | 10 | puppo | Lakeland_terrier | 0.20 |
| 1959 | 10 | puppo | Pembroke | 0.97 |
| 1960 | 10 | puppo | golden_retriever | 0.95 |
| 1961 | 10 | puppo | Eskimo_dog | 0.62 |
| 1962 | 10 | puppo | American_Staffordshire_terrier | 0.93 |
| 1963 | 10 | puppo | chow | 0.81 |
| 1965 | 10 | puppo | German_shepherd | 0.99 |
| 1966 | 10 | puppo | golden_retriever | 0.90 |
| 1967 | 10 | puppo | golden_retriever | 0.25 |
| 1968 | 10 | puppo | Airedale | 0.60 |
| 1970 | 10 | puppo | Labrador_retriever | 0.93 |

| | no_of_likes | retweet_count | \ |
|----|-------------|---------------|---|
| 1 | NaN | 89 | |
| 4 | NaN | 96 | |
| 6 | NaN | 102 | |
| 7 | NaN | 242 | |
| 8 | NaN | 294 | |
| 9 | NaN | 231 | |
| 10 | NaN | 2272 | |
| 11 | NaN | 203 | |
| 13 | NaN | 1198 | |
| 14 | NaN | 576 | |
| 15 | NaN | 72 | |
| 16 | NaN | 568 | |
| 17 | NaN | 138 | |
| 20 | NaN | 162 | |
| 22 | NaN | 61 | |
| 24 | NaN | 358 | |
| 26 | NaN | 264 | |
| 28 | NaN | 258 | |
| 29 | NaN | 282 | |
| 30 | NaN | 211 | |
| 34 | NaN | 105 | |
| 35 | NaN | 83 | |

| | | |
|------|-----|-------|
| 38 | NaN | 102 |
| 40 | NaN | 54 |
| 41 | NaN | 57 |
| 43 | NaN | 134 |
| 44 | NaN | 132 |
| 45 | NaN | 135 |
| 46 | NaN | 2387 |
| 47 | NaN | 18285 |
| 48 | NaN | 66 |
| 49 | NaN | 1241 |
| 51 | NaN | 88 |
| 52 | NaN | 484 |
| 53 | NaN | 58 |
| 54 | NaN | 6618 |
| 56 | NaN | 115 |
| ... | ... | ... |
| 1925 | NaN | 532 |
| 1926 | NaN | 252 |
| 1928 | NaN | 4124 |
| 1929 | NaN | 8952 |
| 1934 | NaN | 894 |
| 1935 | NaN | 1667 |
| 1938 | NaN | 568 |
| 1939 | NaN | 998 |
| 1940 | NaN | 754 |
| 1941 | NaN | 2444 |
| 1942 | NaN | 1069 |
| 1943 | NaN | 2173 |
| 1944 | NaN | 523 |
| 1945 | NaN | 456 |
| 1946 | NaN | 4308 |
| 1947 | NaN | 401 |
| 1948 | NaN | 4319 |
| 1949 | NaN | 11007 |
| 1950 | NaN | 6547 |
| 1951 | NaN | 3040 |
| 1952 | NaN | 3917 |
| 1953 | NaN | 1366 |
| 1954 | NaN | 3220 |
| 1955 | NaN | 19669 |
| 1956 | NaN | 1483 |
| 1957 | NaN | 716 |
| 1958 | NaN | 48265 |
| 1959 | NaN | 8538 |
| 1960 | NaN | 2321 |
| 1961 | NaN | 2697 |
| 1962 | NaN | 1369 |
| 1963 | NaN | 1759 |

| | | |
|------|-----|------|
| 1965 | NaN | 3240 |
| 1966 | NaN | 1704 |
| 1967 | NaN | 5496 |
| 1968 | NaN | 1738 |
| 1970 | NaN | 1480 |

| | tweet_url \ |
|------|---|
| 1 | https://twitter.com/dog_rates/status/667435689... |
| 4 | https://twitter.com/dog_rates/status/667453023... |
| 6 | https://twitter.com/dog_rates/status/667470559... |
| 7 | https://twitter.com/dog_rates/status/667491009... |
| 8 | https://twitter.com/dog_rates/status/667495797... |
| 9 | https://twitter.com/dog_rates/status/667502640... |
| 10 | https://twitter.com/dog_rates/status/667509364... |
| 11 | https://twitter.com/dog_rates/status/667517642... |
| 13 | https://twitter.com/dog_rates/status/667524857... |
| 14 | https://twitter.com/dog_rates/status/667534815... |
| 15 | https://twitter.com/dog_rates/status/667538891... |
| 16 | https://twitter.com/dog_rates/status/667544320... |
| 17 | https://twitter.com/dog_rates/status/667546741... |
| 20 | https://twitter.com/dog_rates/status/667728196... |
| 22 | https://twitter.com/dog_rates/status/667773195... |
| 24 | https://twitter.com/dog_rates/status/667793409... |
| 26 | https://twitter.com/dog_rates/status/667530908... |
| 28 | https://twitter.com/dog_rates/status/667211855... |
| 29 | https://twitter.com/dog_rates/status/667200525... |
| 30 | https://twitter.com/dog_rates/status/666781792... |
| 34 | https://twitter.com/dog_rates/status/666826780... |
| 35 | https://twitter.com/dog_rates/status/666835007... |
| 38 | https://twitter.com/dog_rates/status/666996132... |
| 40 | https://twitter.com/dog_rates/status/667044094... |
| 41 | https://twitter.com/dog_rates/status/667062181... |
| 43 | https://twitter.com/dog_rates/status/667073648... |
| 44 | https://twitter.com/dog_rates/status/667090893... |
| 45 | https://twitter.com/dog_rates/status/667119796... |
| 46 | https://twitter.com/dog_rates/status/667138269... |
| 47 | https://twitter.com/dog_rates/status/667152164... |
| 48 | https://twitter.com/dog_rates/status/667160273... |
| 49 | https://twitter.com/dog_rates/status/667165590... |
| 51 | https://twitter.com/dog_rates/status/667174963... |
| 52 | https://twitter.com/dog_rates/status/667176164... |
| 53 | https://twitter.com/dog_rates/status/667177989... |
| 54 | https://twitter.com/dog_rates/status/667182792... |
| 56 | https://twitter.com/dog_rates/status/667192066... |
| ... | ... |
| 1925 | https://twitter.com/dog_rates/status/674774481... |
| 1926 | https://twitter.com/dog_rates/status/675006312... |
| 1928 | https://twitter.com/dog_rates/status/692752401... |

1929 https://twitter.com/dog_rates/status/837820167...
 1934 https://twitter.com/dog_rates/status/693629975...
 1935 <https://www.gofundme.com/help-my-baby-sierra-g...>
 1938 https://twitter.com/dog_rates/status/693262851...
 1939 https://twitter.com/dog_rates/status/692828166...
 1940 https://twitter.com/dog_rates/status/700151421...
 1941 https://twitter.com/dog_rates/status/747963614...
 1942 https://twitter.com/dog_rates/status/673956914...
 1943 https://twitter.com/dog_rates/status/759793422...
 1944 https://twitter.com/dog_rates/status/693095443...
 1945 https://twitter.com/dog_rates/status/674038233...
 1946 https://twitter.com/dog_rates/status/874296783...
 1947 https://twitter.com/dog_rates/status/673919437...
 1948 https://twitter.com/dog_rates/status/878776093...
 1949 https://twitter.com/dog_rates/status/874012996...
 1950 https://twitter.com/dog_rates/status/793195938...
 1951 https://twitter.com/dog_rates/status/802239329...
 1952 https://twitter.com/dog_rates/status/752519690...
 1953 https://twitter.com/dog_rates/status/743253157...
 1954 https://twitter.com/dog_rates/status/803773340...
 1955 https://twitter.com/dog_rates/status/825535076...
 1956 <https://www.gofundme.com/my-puppys-double-cata...>
 1957 https://twitter.com/dog_rates/status/744995568...
 1958 https://twitter.com/dog_rates/status/822872901...
 1959 https://twitter.com/dog_rates/status/889665388...
 1960 https://twitter.com/dog_rates/status/889531135...
 1961 https://twitter.com/dog_rates/status/867421006...
 1962 <https://www.gofundme.com/servicedogoliver,http...>
 1963 https://twitter.com/dog_rates/status/738537504...
 1965 https://twitter.com/dog_rates/status/787717603...
 1966 https://twitter.com/dog_rates/status/859607811...
 1967 https://twitter.com/dog_rates/status/790946055...
 1968 https://twitter.com/dog_rates/status/756275833...
 1970 https://twitter.com/dog_rates/status/751132876...

image_link

1 <https://pbs.twimg.com/media/CUM10HCW4AEgGSi.jpg>
 4 https://pbs.twimg.com/media/CUNE_OSUwAAAdHhX.jpg
 6 <https://pbs.twimg.com/media/CUNU78YWEAECompB.jpg>
 7 <https://pbs.twimg.com/media/CUNniSlUYAEj1Jl.jpg>
 8 <https://pbs.twimg.com/media/CUNr4-7UwAAg2lq.jpg>
 9 <https://pbs.twimg.com/media/CUNyHTMUyAAQVch.jpg>
 10 <https://pbs.twimg.com/media/CUN40r5UAAAAa5K4.jpg>
 11 https://pbs.twimg.com/media/CUN_wiBUkAAakT0.jpg
 13 https://pbs.twimg.com/media/CUOGUfJW4AA_eni.jpg
 14 https://pbs.twimg.com/media/CUOPYI5UcAAj_n0.jpg
 15 <https://pbs.twimg.com/media/CUOTFZOW4AABsfW.jpg>
 16 <https://pbs.twimg.com/media/CUOYBbbWIAAXQGU.jpg>

17 https://pbs.twimg.com/media/CU0aOWXWcAAO_Jy.jpg
 20 https://pbs.twimg.com/media/CUQ_QahUAAAVQjn.jpg
 22 <https://pbs.twimg.com/media/CURoLrOVEAAaWdR.jpg>
 24 <https://pbs.twimg.com/media/CUR6jqVWsAEgGot.jpg>
 26 <https://pbs.twimg.com/media/CUOL0uGUkAAx7yh.jpg>
 28 <https://pbs.twimg.com/media/CUJppKJWoAA75NP.jpg>
 29 <https://pbs.twimg.com/media/CUJfVMPXIAAgbue.jpg>
 30 https://pbs.twimg.com/media/CUDigRXXIAATI_H.jpg
 34 <https://pbs.twimg.com/media/CUELa0NUkAAscGC.jpg>
 35 <https://pbs.twimg.com/media/CUES51dXIAEahyG.jpg>
 38 <https://pbs.twimg.com/media/CUGlb6iUwAITEbW.jpg>
 40 <https://pbs.twimg.com/media/CUHREBXXAAE6A9b.jpg>
 41 <https://pbs.twimg.com/media/CUHhgvHUA4aB0.jpg>
 43 <https://pbs.twimg.com/media/CUHR8WbWEAEBPgf.jpg>
 44 <https://pbs.twimg.com/media/CUH7oLuUsAELWib.jpg>
 45 <https://pbs.twimg.com/media/CUIV6F7XIAA1tAM.jpg>
 46 <https://pbs.twimg.com/media/CUImtzEVAAAZNJo.jpg>
 47 https://pbs.twimg.com/media/CUIzWk_UwAAfUNq.jpg
 48 <https://pbs.twimg.com/media/CUI6uuaW4AAvCIs.jpg>
 49 https://pbs.twimg.com/media/CUI_kHBWsAAAef5.jpg
 51 <https://pbs.twimg.com/media/CUJIFoJWsAAL3Dc.jpg>
 52 <https://pbs.twimg.com/media/CUJjLtWwsAE-go5.jpg>
 53 <https://pbs.twimg.com/media/CUJK18UWEAEg7AR.jpg>
 54 <https://pbs.twimg.com/media/CUJPNjOWsAAZRqP.jpg>
 56 <https://pbs.twimg.com/media/CUJXpRBXIAAN0yz.jpg>

 1925 <https://pbs.twimg.com/media/CV1HztsWoAAuZwo.jpg>
 1926 <https://pbs.twimg.com/media/CV4aqCwWsAIi30P.jpg>
 1928 https://pbs.twimg.com/tweet_video_thumb/CZ0mhd...
 1929 https://pbs.twimg.com/media/C6CI_jbVAAA3-a1.jpg
 1934 <https://pbs.twimg.com/media/CaBEx3SWEAILZpi.jpg>
 1935 <https://pbs.twimg.com/media/DB5HTBGXUAEOTiK.jpg>
 1938 <https://pbs.twimg.com/media/CZ724fdUYAAytS-.jpg>
 1939 <https://pbs.twimg.com/media/CZ1riVOWwAATfGf.jpg>
 1940 <https://pbs.twimg.com/media/CbdwATgWwAABGID.jpg>
 1941 <https://pbs.twimg.com/media/CmFM7ngXEAIEitfh.jpg>
 1942 <https://pbs.twimg.com/media/CVpgPGwWoAEV7gG.jpg>
 1943 <https://pbs.twimg.com/media/CotUFZEWcAA2Pku.jpg>
 1944 <https://pbs.twimg.com/media/CZ5entwWYAAocEg.jpg>
 1945 https://pbs.twimg.com/media/CVqqMtiVEAEye_L.jpg
 1946 <https://pbs.twimg.com/media/DCIGsROXgAANE0Y.jpg>
 1947 <https://pbs.twimg.com/media/CVo-JuMWwAAet6F.jpg>
 1948 <https://pbs.twimg.com/media/DDIKMXzWOAEibje.jpg>
 1949 <https://pbs.twimg.com/media/DCEeLxjXsAAvNSM.jpg>
 1950 https://pbs.twimg.com/media/CwH_foYWgAEvTyI.jpg
 1951 <https://pbs.twimg.com/media/CyIgaTEVEAA-9zS.jpg>
 1952 <https://pbs.twimg.com/media/CnF8qVDWYAAhOg1.jpg>
 1953 <https://pbs.twimg.com/media/C1CQzFUUYAA5vAu.jpg>

```

1954 https://pbs.twimg.com/media/CyeTku-XcAALkBd.jpg
1955 https://pbs.twimg.com/media/C3TjvitXAAAI-QH.jpg
1956 https://pbs.twimg.com/media/C3MVTehWcAAGNfx.jpg
1957 https://pbs.twimg.com/media/C1bBg4WWEAMjwJu.jpg
1958 https://pbs.twimg.com/media/C2tugXLXgAArJ04.jpg
1959 https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg
1960 https://pbs.twimg.com/media/DFg_2PVWOAEHN3p.jpg
1961 https://pbs.twimg.com/media/DAmmy8FXyaiH8Ty.jpg
1962 https://pbs.twimg.com/media/C2EONHNWQAUWxkP.jpg
1963 https://pbs.twimg.com/media/Cj_P7rSUGAAYQbz.jpg
1965 https://pbs.twimg.com/media/Cu6I9vvWIAAZG0a.jpg
1966 https://pbs.twimg.com/media/C-3wvtxXcAUTuBE.jpg
1967 https://pbs.twimg.com/media/CvoBPWRWgAA4het.jpg
1968 https://pbs.twimg.com/media/Cn7U2xlW8AI9Pqp.jpg
1970 https://pbs.twimg.com/media/CmyPXNOW8AEtaJ-.jpg

```

```
[1409 rows x 14 columns]
```

```
In [131]: df_filtiration = df_breedsmorethan10ratings[df_breedsmorethan10ratings['dog_score'] <=
```

```
In [132]: df_filtiration.groupby('dog_breed')['dog_score'].describe()
```

```
Out[132]:
```

| | count | mean | std | min | 25% | 50% | \ |
|--------------------------------|-------|-----------|----------|------|-------|------|---|
| dog_breed | | | | | | | |
| Airedale | 12.0 | 9.833333 | 2.249579 | 6.0 | 9.25 | 10.5 | |
| American_Staffordshire_terrier | 16.0 | 10.312500 | 3.070695 | 0.0 | 10.00 | 11.0 | |
| Bernese_mountain_dog | 11.0 | 11.272727 | 1.103713 | 10.0 | 10.00 | 12.0 | |
| Blenheim_spaniel | 10.0 | 11.100000 | 1.100505 | 9.0 | 10.25 | 11.5 | |
| Border_collie | 12.0 | 10.583333 | 2.503028 | 4.0 | 9.75 | 11.5 | |
| Boston_bull | 12.0 | 9.416667 | 2.108784 | 5.0 | 8.75 | 10.0 | |
| Cardigan | 21.0 | 11.142857 | 1.590148 | 7.0 | 10.00 | 11.0 | |
| Chesapeake_Bay_retriever | 31.0 | 10.838710 | 1.439833 | 8.0 | 10.00 | 10.0 | |
| Chihuahua | 90.0 | 10.500000 | 2.089299 | 3.0 | 10.00 | 11.0 | |
| English_springer | 10.0 | 10.500000 | 2.223611 | 5.0 | 10.00 | 10.5 | |
| Eskimo_dog | 22.0 | 11.409091 | 1.402688 | 9.0 | 10.00 | 12.0 | |
| French_bulldog | 31.0 | 11.193548 | 1.796652 | 8.0 | 10.00 | 12.0 | |
| German_shepherd | 21.0 | 11.000000 | 1.449138 | 8.0 | 10.00 | 11.0 | |
| Great_Dane | 11.0 | 10.818182 | 1.940009 | 7.0 | 10.00 | 11.0 | |
| Great_Pyrenees | 14.0 | 11.428571 | 1.222500 | 9.0 | 11.00 | 11.5 | |
| Italian_greyhound | 17.0 | 10.000000 | 1.870829 | 6.0 | 8.00 | 11.0 | |
| Labrador_retriever | 103.0 | 11.203883 | 1.345721 | 7.0 | 10.00 | 11.0 | |
| Lakeland_terrier | 17.0 | 10.235294 | 1.751050 | 7.0 | 9.00 | 10.0 | |
| Maltese_dog | 19.0 | 9.736842 | 2.077448 | 2.0 | 9.50 | 10.0 | |
| Norwegian_elkhound | 11.0 | 10.909091 | 1.445998 | 8.0 | 10.50 | 11.0 | |
| Old_English_sheepdog | 14.0 | 10.857143 | 1.511858 | 8.0 | 10.00 | 11.0 | |
| Pekinese | 14.0 | 11.000000 | 1.921538 | 7.0 | 10.00 | 12.0 | |
| Pembroke | 94.0 | 11.425532 | 1.719523 | 4.0 | 11.00 | 12.0 | |
| Pomeranian | 40.0 | 10.975000 | 1.404891 | 6.0 | 10.00 | 11.0 | |

| | | | | | | |
|-----------------------------|-------|-----------|----------|------|-------|------|
| Rottweiler | 18.0 | 11.000000 | 1.748949 | 7.0 | 10.00 | 11.5 |
| Samoyed | 42.0 | 11.690476 | 1.352290 | 7.0 | 11.00 | 12.0 |
| Shetland_sheepdog | 18.0 | 10.388889 | 1.719743 | 7.0 | 9.25 | 10.5 |
| Shih-Tzu | 20.0 | 10.350000 | 1.308877 | 8.0 | 9.75 | 10.5 |
| Siberian_husky | 20.0 | 11.300000 | 1.128576 | 9.0 | 10.75 | 11.0 |
| Staffordshire_bullterrier | 20.0 | 10.750000 | 1.409554 | 8.0 | 10.00 | 11.0 |
| West_Highland_white_terrier | 15.0 | 10.066667 | 1.486447 | 8.0 | 9.50 | 10.0 |
| Yorkshire_terrier | 10.0 | 10.200000 | 1.686548 | 8.0 | 9.00 | 9.5 |
| basset | 17.0 | 10.705882 | 2.201270 | 5.0 | 10.00 | 11.0 |
| beagle | 20.0 | 10.150000 | 1.531253 | 6.0 | 9.75 | 10.0 |
| borzoi | 9.0 | 10.111111 | 2.204793 | 7.0 | 9.00 | 10.0 |
| boxer | 11.0 | 10.545455 | 1.572491 | 8.0 | 10.00 | 10.0 |
| chow | 47.0 | 11.404255 | 1.361897 | 7.0 | 11.00 | 12.0 |
| cocker_spaniel | 30.0 | 11.333333 | 1.184187 | 9.0 | 10.25 | 11.5 |
| collie | 11.0 | 10.909091 | 1.221028 | 9.0 | 10.00 | 11.0 |
| dalmatian | 12.0 | 9.333333 | 2.188122 | 5.0 | 8.00 | 9.5 |
| golden_retriever | 155.0 | 11.561290 | 1.324513 | 5.0 | 11.00 | 12.0 |
| kelpie | 13.0 | 11.307692 | 1.109400 | 10.0 | 11.00 | 11.0 |
| kuvasz | 16.0 | 10.625000 | 2.276694 | 4.0 | 10.00 | 11.5 |
| malamute | 33.0 | 10.878788 | 1.218544 | 8.0 | 10.00 | 11.0 |
| malinois | 11.0 | 10.818182 | 1.990888 | 6.0 | 10.00 | 11.0 |
| miniature_pinscher | 24.0 | 10.041667 | 2.612331 | 2.0 | 9.75 | 11.0 |
| pug | 62.0 | 10.241935 | 1.816910 | 3.0 | 9.25 | 10.0 |
| schipperke | 13.0 | 10.923077 | 1.382120 | 8.0 | 10.00 | 11.0 |
| soft-coated_wheaten_terrier | 12.0 | 8.166667 | 3.128559 | 3.0 | 5.50 | 9.5 |
| standard_poodle | 11.0 | 10.727273 | 1.678744 | 9.0 | 9.50 | 10.0 |
| toy_poodle | 50.0 | 11.000000 | 1.245400 | 7.0 | 10.00 | 11.0 |
| vizsla | 13.0 | 10.846154 | 1.281025 | 8.0 | 11.00 | 11.0 |
| whippet | 11.0 | 10.181818 | 1.834022 | 8.0 | 8.50 | 10.0 |

75% max

| | | |
|--------------------------------|-------|------|
| dog_breed | | |
| Airedale | 11.25 | 12.0 |
| American_Staffordshire_terrier | 12.00 | 13.0 |
| Bernese_mountain_dog | 12.00 | 13.0 |
| Blenheim_spaniel | 12.00 | 12.0 |
| Border_collie | 12.00 | 13.0 |
| Boston_bull | 11.00 | 12.0 |
| Cardigan | 12.00 | 13.0 |
| Chesapeake_Bay_retriever | 12.00 | 13.0 |
| Chihuahua | 12.00 | 14.0 |
| English_springer | 12.00 | 13.0 |
| Eskimo_dog | 12.00 | 14.0 |
| French_bulldog | 12.00 | 14.0 |
| German_shepherd | 12.00 | 13.0 |
| Great_Dane | 12.00 | 13.0 |
| Great_Pyrenees | 12.00 | 13.0 |
| Italian_greyhound | 11.00 | 13.0 |

| | | |
|-----------------------------|-------|------|
| Labrador_retriever | 12.00 | 14.0 |
| Lakeland_terrier | 11.00 | 13.0 |
| Maltese_dog | 10.50 | 12.0 |
| Norwegian_elkhound | 12.00 | 13.0 |
| Old_English_sheepdog | 12.00 | 13.0 |
| Pekinese | 12.00 | 13.0 |
| Pembroke | 12.00 | 14.0 |
| Pomeranian | 12.00 | 14.0 |
| Rottweiler | 12.00 | 14.0 |
| Samoyed | 13.00 | 14.0 |
| Shetland_sheepdog | 11.75 | 13.0 |
| Shih-Tzu | 11.00 | 13.0 |
| Siberian_husky | 12.00 | 13.0 |
| Staffordshire_bullterrier | 12.00 | 13.0 |
| West_Highland_white_terrier | 10.50 | 13.0 |
| Yorkshire_terrier | 11.75 | 13.0 |
| basset | 12.00 | 13.0 |
| beagle | 11.00 | 13.0 |
| borzoi | 11.00 | 13.0 |
| boxer | 11.50 | 13.0 |
| chow | 12.00 | 13.0 |
| cocker_spaniel | 12.00 | 13.0 |
| collie | 12.00 | 13.0 |
| dalmatian | 10.25 | 13.0 |
| golden_retriever | 12.00 | 14.0 |
| kelpie | 12.00 | 13.0 |
| kuvasz | 12.00 | 13.0 |
| malamute | 12.00 | 13.0 |
| malinois | 12.00 | 13.0 |
| miniature_pinscher | 12.00 | 12.0 |
| pug | 11.00 | 13.0 |
| schipperke | 12.00 | 13.0 |
| soft-coated_wheaten_terrier | 10.25 | 12.0 |
| standard_poodle | 11.50 | 14.0 |
| toy_poodle | 12.00 | 13.0 |
| vizsla | 11.00 | 13.0 |
| whippet | 12.00 | 13.0 |

In [133]: df_filtiration.groupby('dog_breed')['dog_score'].mean().sort_values(ascending=False).n

Out[133]: dog_breed

| | |
|------------------|-----------|
| Samoyed | 11.690476 |
| golden_retriever | 11.561290 |
| Great_Pyrenees | 11.428571 |
| Pembroke | 11.425532 |
| Eskimo_dog | 11.409091 |
| chow | 11.404255 |
| cocker_spaniel | 11.333333 |

```

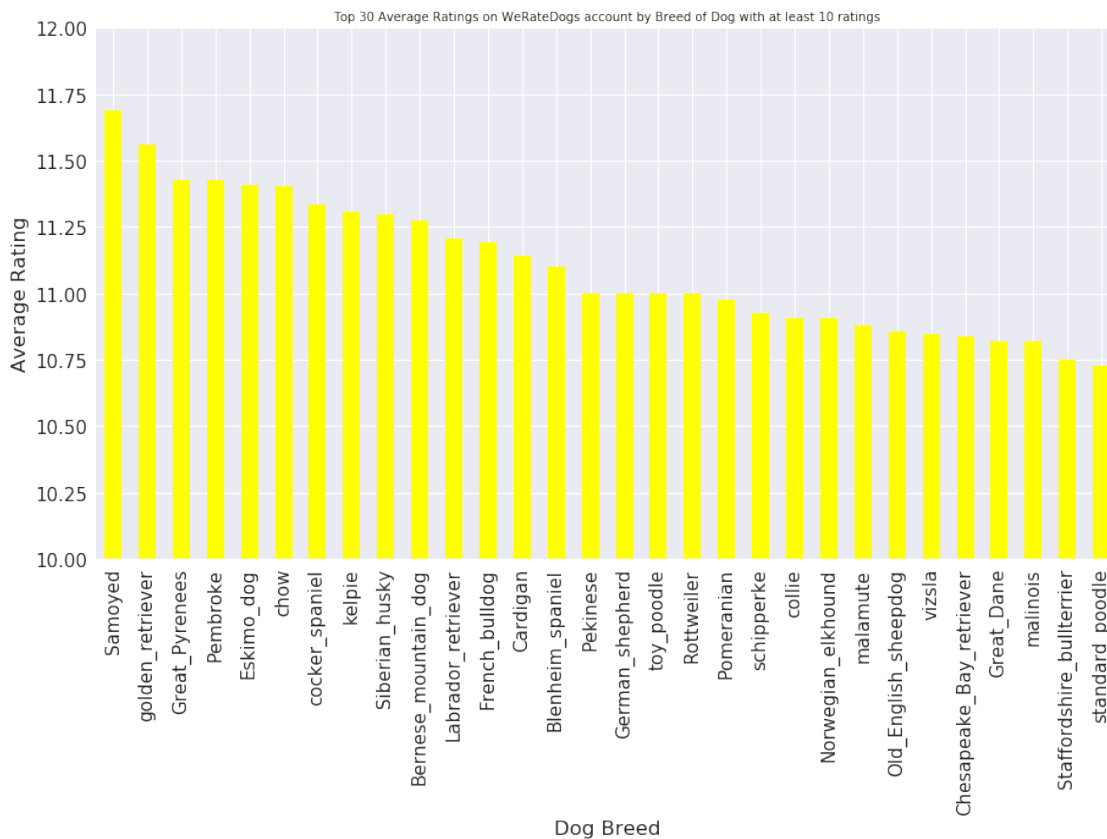
kelpie          11.307692
Siberian_husky  11.300000
Bernese_mountain_dog  11.272727
Name: dog_score, dtype: float64

```

```

In [135]: fig = plt.figure(figsize=(15,8))
df_filtiration.groupby('dog_breed')['dog_score'].mean().sort_values(ascending=False).n
plt.title("Top 30 Average Ratings on WeRateDogs account by Breed of Dog with at least
plt.ylabel("Average Rating")
plt.xlabel("Dog Breed")
plt.ylim(10,12);

```



```

In [136]: df_filtiration.head()

```

```

Out[136]:
   tweet_id  tweet_date  tweet_source \
1  667435689202614272  2015-11-19 20:14:03  Twitter for iPhone
4  667453023279554560  2015-11-19 21:22:56  Twitter Web Client
6  667470559035432960  2015-11-19 22:32:36  Twitter Web Client
7  667491009379606528  2015-11-19 23:53:52  Twitter Web Client
8  667495797102141441  2015-11-20 00:12:54  Twitter Web Client

```

| | tweet_text | dog_name | dog_score | \ |
|---|---|----------|-----------|---|
| 1 | Ermergerd 12/10 https://t.co/PQni2sjPsm | None | 12 | |
| 4 | Meet Cupcake. I would do unspeakable things fo... | Cupcake | 11 | |
| 6 | This is a northern Wahoo named Kohl. He runs t... | None | 11 | |
| 7 | Two dogs in this one. Both are rare Jujitsu Py... | None | 7 | |
| 8 | This is Philippe from Soviet Russia. Commandin... | Philippe | 9 | |

| | out_of | dog_type | dog_breed | probability_AI_calc | no_of_likes | \ |
|---|--------|----------|--------------------|---------------------|-------------|---|
| 1 | 10 | None | Rottweiler | 1.00 | NaN | |
| 4 | 10 | None | Labrador_retriever | 0.83 | NaN | |
| 6 | 10 | None | toy_poodle | 0.30 | NaN | |
| 7 | 10 | None | borzoi | 0.85 | NaN | |
| 8 | 10 | None | Chihuahua | 0.14 | NaN | |

| | retweet_count | tweet_url | \ |
|---|---------------|---|---|
| 1 | 89 | https://twitter.com/dog_rates/status/667435689... | |
| 4 | 96 | https://twitter.com/dog_rates/status/667453023... | |
| 6 | 102 | https://twitter.com/dog_rates/status/667470559... | |
| 7 | 242 | https://twitter.com/dog_rates/status/667491009... | |
| 8 | 294 | https://twitter.com/dog_rates/status/667495797... | |

| | image_link |
|---|---|
| 1 | https://pbs.twimg.com/media/CUM10HCW4AEgGSi.jpg |
| 4 | https://pbs.twimg.com/media/CUNE_OSUwAAhHhX.jpg |
| 6 | https://pbs.twimg.com/media/CUNU78YWEAECmpB.jpg |
| 7 | https://pbs.twimg.com/media/CUNniSlUYAEj1Jl.jpg |
| 8 | https://pbs.twimg.com/media/CUNr4-7UwAAg2lq.jpg |

```
In [137]: df_filtiration['no_of_likes'].sum()
```

```
Out[137]: 0.0
```

```
In [ ]:
```