

# wrangle\_act

July 31, 2018

## 1 Data Wrangling Project

In this project I will wrangle and analyze the tweet archive of Twitter user @dog\_rates. @dog\_rates that rates people's dogs with a some tweet about the dog. These ratings almost always have a denominator of 10. @dog\_rates asks people to send photos of their dogs which are rated on a scale of one to ten, but are invariably given ratings in excess of the maximum, such as "13/10".

### 1.1 DataSet

- The First Dataset is a csv file which we need to access it by downloading it
- The Second Dataset is a tweet image predictions, i.e., what breed of dog (or other object, animal, etc.) is present in each tweet according to a neural network. This file (image\_predictions.tsv) is hosted on Udacity's servers and should be downloaded programmatically using the Requests library and the following URL: [https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad\\_image-predictions/image-predictions.tsv](https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad_image-predictions/image-predictions.tsv)
- The Third Dataset is need to be access using twitter API in and save in a file in json format after that it need to be converted to dataframe

```
In [4]: import pandas as pd
import numpy as np
import tweepy
import requests
import json
import os
import re
from datetime import datetime
import matplotlib.pyplot as plt
%matplotlib inline
```

```
In [5]: df=pd.read_csv('twitter-archive-enhanced.csv')
df.head()
```

```
Out[5]:
```

	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

	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

	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...

	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

	retweeted_status_user_id	retweeted_status_timestamp \
0	NaN	NaN
1	NaN	NaN
2	NaN	NaN
3	NaN	NaN
4	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

	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

```
In [6]: #import the tweet image predictions using requests library#import
url = 'https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad_image-predictio
```

```

response = requests.get(url)

with open('image-predictions.tsv', mode = 'wb') as file:
    file.write(response.content)

#read the image predictions file
df2 = pd.read_csv('image-predictions.tsv', sep = '\t')

In [7]: consumer_key = 'nUEeH00Uee42DaNH713EX1wv'
        consumer_secret = 'Y232RMHyf1YCsjPLZVOLmdHrD0942AjjPHcIibYUwrtzBORkFs'
        access_token = '988144052439838722-qv8g4xTxhB1LcytyAk94oUukgcPjmJ3'
        access_secret = 'iXC81zHUHjj5mx2EDIRHnFYEP1sLbzpjzobzASS73RpBc'

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

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

In [8]: #add tweets to tweet_json.txt
        with open('tweet_json.txt', 'a', encoding='utf8') as f:
            for tweet_id in df['tweet_id']:
                try:
                    tweet = api.get_status(tweet_id, tweet_mode='extended')
                    json.dump(tweet._json, f)
                    f.write('\n')
                except:
                    continue

Rate limit reached. Sleeping for: 735
Rate limit reached. Sleeping for: 736

In [9]: #append the tweets to a list
        tweets_data = []

        tweet_file = open('tweet_json.txt', "r")

        for line in tweet_file:
            try:
                tweet = json.loads(line)
                tweets_data.append(tweet)
            except:
                continue

        tweet_file.close()

In [10]: json_tweets=pd.DataFrame()

In [11]: #add the necessary columns to the data frame
        json_tweets['id'] = list(map(lambda tweet: tweet['id'], tweets_data))

```

```
json_tweets['retweet_count'] = list(map(lambda tweet: tweet['retweet_count'], tweets_data))
json_tweets['favorite_count'] = list(map(lambda tweet: tweet['favorite_count'], tweets_data))
```

## 2 Assess

### 2.1 After gathering each of the above pieces of data, I assessed them visually and programmatically for quality and tidiness issues.

In [12]: df

```
Out[12]:
```

	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...

```

[illegible]

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 <a href="https://t.co/ws2AMaWpPW">https://t.co/ws2AMaWpPW</a>	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	<a href="https://twitter.com/dog_rates/status/892420643...">https://twitter.com/dog_rates/status/892420643...</a>	13
1	<a href="https://twitter.com/dog_rates/status/892177421...">https://twitter.com/dog_rates/status/892177421...</a>	13
2	<a href="https://twitter.com/dog_rates/status/891815181...">https://twitter.com/dog_rates/status/891815181...</a>	12
3	<a href="https://twitter.com/dog_rates/status/891689557...">https://twitter.com/dog_rates/status/891689557...</a>	13
4	<a href="https://twitter.com/dog_rates/status/891327558...">https://twitter.com/dog_rates/status/891327558...</a>	12
5	<a href="https://twitter.com/dog_rates/status/891087950...">https://twitter.com/dog_rates/status/891087950...</a>	13
6	<a href="https://gofundme.com/ydvmve-surgery-for-jax,ht...">https://gofundme.com/ydvmve-surgery-for-jax,ht...</a>	13
7	<a href="https://twitter.com/dog_rates/status/890729181...">https://twitter.com/dog_rates/status/890729181...</a>	13
8	<a href="https://twitter.com/dog_rates/status/890609185...">https://twitter.com/dog_rates/status/890609185...</a>	13
9	<a href="https://twitter.com/dog_rates/status/890240255...">https://twitter.com/dog_rates/status/890240255...</a>	14

10	<a href="https://twitter.com/dog_rates/status/890006608...">https://twitter.com/dog_rates/status/890006608...</a>	13
11	<a href="https://twitter.com/dog_rates/status/889880896...">https://twitter.com/dog_rates/status/889880896...</a>	13
12	<a href="https://twitter.com/dog_rates/status/889665388...">https://twitter.com/dog_rates/status/889665388...</a>	13
13	<a href="https://twitter.com/dog_rates/status/889638837...">https://twitter.com/dog_rates/status/889638837...</a>	12
14	<a href="https://twitter.com/dog_rates/status/889531135...">https://twitter.com/dog_rates/status/889531135...</a>	13
15	<a href="https://twitter.com/dog_rates/status/889278841...">https://twitter.com/dog_rates/status/889278841...</a>	13
16	<a href="https://twitter.com/dog_rates/status/888917238...">https://twitter.com/dog_rates/status/888917238...</a>	12
17	<a href="https://twitter.com/dog_rates/status/888804989...">https://twitter.com/dog_rates/status/888804989...</a>	13
18	<a href="https://twitter.com/dog_rates/status/888554962...">https://twitter.com/dog_rates/status/888554962...</a>	13
19	<a href="https://twitter.com/dog_rates/status/887473957...">https://twitter.com/dog_rates/status/887473957...</a>	13
20	<a href="https://twitter.com/dog_rates/status/888078434...">https://twitter.com/dog_rates/status/888078434...</a>	12
21	<a href="https://twitter.com/dog_rates/status/887705289...">https://twitter.com/dog_rates/status/887705289...</a>	13
22	<a href="https://twitter.com/dog_rates/status/887517139...">https://twitter.com/dog_rates/status/887517139...</a>	14
23	<a href="https://twitter.com/dog_rates/status/887473957...">https://twitter.com/dog_rates/status/887473957...</a>	13
24	<a href="https://twitter.com/dog_rates/status/887343217...">https://twitter.com/dog_rates/status/887343217...</a>	13
25	<a href="https://twitter.com/dog_rates/status/887101392...">https://twitter.com/dog_rates/status/887101392...</a>	12
26	<a href="https://twitter.com/dog_rates/status/886983233...">https://twitter.com/dog_rates/status/886983233...</a>	13
27	<a href="https://www.gofundme.com/mingusneedsus">https://www.gofundme.com/mingusneedsus</a> , <a href="https://twitter.com/dog_rates/status/886680336...">https://twitter.com/dog_rates/status/886680336...</a>	13
28	<a href="https://twitter.com/dog_rates/status/886680336...">https://twitter.com/dog_rates/status/886680336...</a>	13
29	<a href="https://twitter.com/dog_rates/status/886366144...">https://twitter.com/dog_rates/status/886366144...</a>	12
...	...	...
2326	<a href="https://twitter.com/dog_rates/status/666411507...">https://twitter.com/dog_rates/status/666411507...</a>	2
2327	<a href="https://twitter.com/dog_rates/status/666407126...">https://twitter.com/dog_rates/status/666407126...</a>	7
2328	<a href="https://twitter.com/dog_rates/status/666396247...">https://twitter.com/dog_rates/status/666396247...</a>	9
2329	<a href="https://twitter.com/dog_rates/status/666373753...">https://twitter.com/dog_rates/status/666373753...</a>	11
2330	<a href="https://twitter.com/dog_rates/status/666362758...">https://twitter.com/dog_rates/status/666362758...</a>	6
2331	<a href="https://twitter.com/dog_rates/status/666353288...">https://twitter.com/dog_rates/status/666353288...</a>	8
2332	<a href="https://twitter.com/dog_rates/status/666345417...">https://twitter.com/dog_rates/status/666345417...</a>	10
2333	<a href="https://twitter.com/dog_rates/status/666337882...">https://twitter.com/dog_rates/status/666337882...</a>	9
2334	<a href="https://twitter.com/dog_rates/status/666293911...">https://twitter.com/dog_rates/status/666293911...</a>	3
2335	<a href="https://twitter.com/dog_rates/status/666287406...">https://twitter.com/dog_rates/status/666287406...</a>	1
2336	<a href="https://twitter.com/dog_rates/status/666273097...">https://twitter.com/dog_rates/status/666273097...</a>	11
2337	<a href="https://twitter.com/dog_rates/status/666268910...">https://twitter.com/dog_rates/status/666268910...</a>	10
2338	<a href="https://twitter.com/dog_rates/status/666104133...">https://twitter.com/dog_rates/status/666104133...</a>	1
2339	<a href="https://twitter.com/dog_rates/status/666102155...">https://twitter.com/dog_rates/status/666102155...</a>	11
2340	<a href="https://twitter.com/dog_rates/status/666099513...">https://twitter.com/dog_rates/status/666099513...</a>	8
2341	<a href="https://twitter.com/dog_rates/status/666094000...">https://twitter.com/dog_rates/status/666094000...</a>	9
2342	<a href="https://twitter.com/dog_rates/status/666082916...">https://twitter.com/dog_rates/status/666082916...</a>	6
2343	<a href="https://twitter.com/dog_rates/status/666073100...">https://twitter.com/dog_rates/status/666073100...</a>	10
2344	<a href="https://twitter.com/dog_rates/status/666071193...">https://twitter.com/dog_rates/status/666071193...</a>	9
2345	<a href="https://twitter.com/dog_rates/status/666063827...">https://twitter.com/dog_rates/status/666063827...</a>	10
2346	<a href="https://twitter.com/dog_rates/status/666058600...">https://twitter.com/dog_rates/status/666058600...</a>	8
2347	<a href="https://twitter.com/dog_rates/status/666057090...">https://twitter.com/dog_rates/status/666057090...</a>	9
2348	<a href="https://twitter.com/dog_rates/status/666055525...">https://twitter.com/dog_rates/status/666055525...</a>	10
2349	<a href="https://twitter.com/dog_rates/status/666051853...">https://twitter.com/dog_rates/status/666051853...</a>	2
2350	<a href="https://twitter.com/dog_rates/status/666050758...">https://twitter.com/dog_rates/status/666050758...</a>	10
2351	<a href="https://twitter.com/dog_rates/status/666049248...">https://twitter.com/dog_rates/status/666049248...</a>	5
2352	<a href="https://twitter.com/dog_rates/status/666044226...">https://twitter.com/dog_rates/status/666044226...</a>	6

2353	<a href="https://twitter.com/dog_rates/status/666033412...">https://twitter.com/dog_rates/status/666033412...</a>	9
2354	<a href="https://twitter.com/dog_rates/status/666029285...">https://twitter.com/dog_rates/status/666029285...</a>	7
2355	<a href="https://twitter.com/dog_rates/status/666020888...">https://twitter.com/dog_rates/status/666020888...</a>	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 [13]: df2

Out [13]:	tweet_id	jpg_url \
0	666020888022790149	<a href="https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg">https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg</a>
1	666029285002620928	<a href="https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg">https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg</a>
2	666033412701032449	<a href="https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg">https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg</a>
3	666044226329800704	<a href="https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg">https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg</a>
4	666049248165822465	<a href="https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg">https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg</a>
5	666050758794694657	<a href="https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg">https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg</a>
6	666051853826850816	<a href="https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg">https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg</a>
7	666055525042405380	<a href="https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg">https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg</a>
8	666057090499244032	<a href="https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg">https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg</a>
9	666058600524156928	<a href="https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg">https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg</a>
10	666063827256086533	<a href="https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg">https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg</a>
11	666071193221509120	<a href="https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg">https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg</a>
12	666073100786774016	<a href="https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg">https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg</a>
13	666082916733198337	<a href="https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg">https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg</a>
14	666094000022159362	<a href="https://pbs.twimg.com/media/CT5w9gUW4AAaBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAaBNN.jpg</a>
15	666099513787052032	<a href="https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg">https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg</a>
16	666102155909144576	<a href="https://pbs.twimg.com/media/CT54YGiwUAENZnoK.jpg">https://pbs.twimg.com/media/CT54YGiwUAENZnoK.jpg</a>
17	666104133288665088	<a href="https://pbs.twimg.com/media/CT56LSZWAA1Jj2.jpg">https://pbs.twimg.com/media/CT56LSZWAA1Jj2.jpg</a>
18	666268910803644416	<a href="https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg">https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg</a>
19	666273097616637952	<a href="https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg">https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg</a>
20	666287406224695296	<a href="https://pbs.twimg.com/media/CT8g3BpUEAAuFjg.jpg">https://pbs.twimg.com/media/CT8g3BpUEAAuFjg.jpg</a>
21	666293911632134144	<a href="https://pbs.twimg.com/media/CT8mx7KW4AEQu8N.jpg">https://pbs.twimg.com/media/CT8mx7KW4AEQu8N.jpg</a>
22	666337882303524864	<a href="https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg">https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg</a>
23	666345417576210432	<a href="https://pbs.twimg.com/media/CT9Vn7PWAAA_ZCM.jpg">https://pbs.twimg.com/media/CT9Vn7PWAAA_ZCM.jpg</a>

24	666353288456101888	<a href="https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg">https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg</a>
25	666362758909284353	<a href="https://pbs.twimg.com/media/CT9lXGsUcAAyUft.jpg">https://pbs.twimg.com/media/CT9lXGsUcAAyUft.jpg</a>
26	666373753744588802	<a href="https://pbs.twimg.com/media/CT9vZEYWUAA1Z05.jpg">https://pbs.twimg.com/media/CT9vZEYWUAA1Z05.jpg</a>
27	666396247373291520	<a href="https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg">https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg</a>
28	666407126856765440	<a href="https://pbs.twimg.com/media/CT-NvwmW4AAugGZ.jpg">https://pbs.twimg.com/media/CT-NvwmW4AAugGZ.jpg</a>
29	666411507551481857	<a href="https://pbs.twimg.com/media/CT-RugiWIAELEaq.jpg">https://pbs.twimg.com/media/CT-RugiWIAELEaq.jpg</a>
...	...	...
2045	886366144734445568	<a href="https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg">https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg</a>
2046	886680336477933568	<a href="https://pbs.twimg.com/media/DE4fEDzWAAAYHMM.jpg">https://pbs.twimg.com/media/DE4fEDzWAAAYHMM.jpg</a>
2047	886736880519319552	<a href="https://pbs.twimg.com/media/DE5Se8FXcAAJFx4.jpg">https://pbs.twimg.com/media/DE5Se8FXcAAJFx4.jpg</a>
2048	886983233522544640	<a href="https://pbs.twimg.com/media/DE8yicJW0AAAvBJ.jpg">https://pbs.twimg.com/media/DE8yicJW0AAAvBJ.jpg</a>
2049	887101392804085760	<a href="https://pbs.twimg.com/media/DE-eAq6UwAA-jaE.jpg">https://pbs.twimg.com/media/DE-eAq6UwAA-jaE.jpg</a>
2050	887343217045368832	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88734...">https://pbs.twimg.com/ext_tw_video_thumb/88734...</a>
2051	887473957103951883	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2052	887517139158093824	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88751...">https://pbs.twimg.com/ext_tw_video_thumb/88751...</a>
2053	887705289381826560	<a href="https://pbs.twimg.com/media/DFHDQBbXgAEqY7t.jpg">https://pbs.twimg.com/media/DFHDQBbXgAEqY7t.jpg</a>
2054	888078434458587136	<a href="https://pbs.twimg.com/media/DFMWn56WsAAKA7B.jpg">https://pbs.twimg.com/media/DFMWn56WsAAKA7B.jpg</a>
2055	888202515573088257	<a href="https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg">https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg</a>
2056	888554962724278272	<a href="https://pbs.twimg.com/media/DFTH_0-UQAAACu20.jpg">https://pbs.twimg.com/media/DFTH_0-UQAAACu20.jpg</a>
2057	888804989199671297	<a href="https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg">https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg</a>
2058	888917238123831296	<a href="https://pbs.twimg.com/media/DFYRgsOUQAARGh0.jpg">https://pbs.twimg.com/media/DFYRgsOUQAARGh0.jpg</a>
2059	889278841981685760	<a href="https://pbs.twimg.com/ext_tw_video_thumb/88927...">https://pbs.twimg.com/ext_tw_video_thumb/88927...</a>
2060	889531135344209921	<a href="https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg">https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg</a>
2061	889638837579907072	<a href="https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg">https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg</a>
2062	889665388333682689	<a href="https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg">https://pbs.twimg.com/media/DFi579UWsAAatzw.jpg</a>
2063	889880896479866881	<a href="https://pbs.twimg.com/media/DF199B1WsAITKsg.jpg">https://pbs.twimg.com/media/DF199B1WsAITKsg.jpg</a>
2064	890006608113172480	<a href="https://pbs.twimg.com/media/DFnwsY4WAAAMliS.jpg">https://pbs.twimg.com/media/DFnwsY4WAAAMliS.jpg</a>
2065	890240255349198849	<a href="https://pbs.twimg.com/media/DFrEyVuW0AA03t9.jpg">https://pbs.twimg.com/media/DFrEyVuW0AA03t9.jpg</a>
2066	890609185150312448	<a href="https://pbs.twimg.com/media/DFwUU__XcAEpyXI.jpg">https://pbs.twimg.com/media/DFwUU__XcAEpyXI.jpg</a>
2067	890729181411237888	<a href="https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg">https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg</a>
2068	890971913173991426	<a href="https://pbs.twimg.com/media/DF1eOmZXUAAALUcq.jpg">https://pbs.twimg.com/media/DF1eOmZXUAAALUcq.jpg</a>
2069	891087950875897856	<a href="https://pbs.twimg.com/media/DF3HwyEWsAABqE6.jpg">https://pbs.twimg.com/media/DF3HwyEWsAABqE6.jpg</a>
2070	891327558926688256	<a href="https://pbs.twimg.com/media/DF6hr6BUMAAZgZT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAZgZT.jpg</a>
2071	891689557279858688	<a href="https://pbs.twimg.com/media/DF_q7IAWsAEuuN8.jpg">https://pbs.twimg.com/media/DF_q7IAWsAEuuN8.jpg</a>
2072	891815181378084864	<a href="https://pbs.twimg.com/media/DGBdLU1WsAANxJ9.jpg">https://pbs.twimg.com/media/DGBdLU1WsAANxJ9.jpg</a>
2073	892177421306343426	<a href="https://pbs.twimg.com/media/DGGmoV4XsAAUL6n.jpg">https://pbs.twimg.com/media/DGGmoV4XsAAUL6n.jpg</a>
2074	892420643555336193	<a href="https://pbs.twimg.com/media/DGKD1-bXoAAIAUK.jpg">https://pbs.twimg.com/media/DGKD1-bXoAAIAUK.jpg</a>

	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	terrapi
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 [14]: json\_tweets

Out[14]:	id	retweet_count	favorite_count
0	892420643555336193	8532	38588
1	892177421306343426	6270	33086
2	891815181378084864	4156	24912
3	891689557279858688	8660	42009
4	891327558926688256	9411	40140
5	891087950875897856	3115	20128
6	890971913173991426	2073	11797
7	890729181411237888	18907	65209

8	890609185150312448	4270	27667
9	890240255349198849	7427	31787
10	890006608113172480	7341	30543
11	889880896479866881	4976	27651
12	889665388333682689	10063	47906
13	889638837579907072	4548	27046
14	889531135344209921	2236	15030
15	889278841981685760	5419	25167
16	888917238123831296	4501	28948
17	888804989199671297	4339	25460
18	888554962724278272	3580	19809
19	892420643555336193	8532	38588
20	892177421306343426	6270	33086
21	891815181378084864	4156	24912
22	891689557279858688	8660	42009
23	891327558926688256	9411	40140
24	891087950875897856	3115	20128
25	890971913173991426	2073	11797
26	890729181411237888	18906	65209
27	890609185150312448	4270	27667
28	890240255349198849	7427	31787
29	890006608113172480	7341	30543
...	...	...	...
31189	666411507551481857	327	446
31190	666407126856765440	40	109
31191	666396247373291520	84	168
31192	666373753744588802	93	189
31193	666362758909284353	569	777
31194	666353288456101888	72	220
31195	666345417576210432	136	296
31196	666337882303524864	91	198
31197	666293911632134144	354	505
31198	666287406224695296	64	149
31199	666273097616637952	76	174
31200	666268910803644416	35	103
31201	666104133288665088	6592	14307
31202	666102155909144576	12	80
31203	666099513787052032	67	155
31204	666094000022159362	73	164
31205	666082916733198337	44	119
31206	666073100786774016	163	322
31207	666071193221509120	60	148
31208	666063827256086533	217	476
31209	666058600524156928	57	111
31210	666057090499244032	141	298
31211	666055525042405380	249	434
31212	666051853826850816	848	1214
31213	666050758794694657	58	133

31214	666049248165822465	40	109
31215	666044226329800704	141	299
31216	666033412701032449	44	125
31217	666029285002620928	47	130
31218	666020888022790149	515	2563

[31219 rows x 3 columns]

In [15]: df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2356 entries, 0 to 2355
Data columns (total 17 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
dtypes: float64(4), int64(3), object(10)
memory usage: 313.0+ KB
```

In [16]: df2.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2075 entries, 0 to 2074
Data columns (total 12 columns):
tweet_id    2075 non-null int64
jpg_url     2075 non-null object
img_num     2075 non-null int64
p1          2075 non-null object
p1_conf     2075 non-null float64
p1_dog      2075 non-null bool
p2          2075 non-null object
p2_conf     2075 non-null float64
p2_dog      2075 non-null bool
p3          2075 non-null object
```

```
p3_conf      2075 non-null float64
p3_dog       2075 non-null bool
dtypes: bool(3), float64(3), int64(2), object(4)
memory usage: 152.1+ KB
```

```
In [17]: json_tweets.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 31219 entries, 0 to 31218
Data columns (total 3 columns):
id              31219 non-null int64
retweet_count   31219 non-null int64
favorite_count  31219 non-null int64
dtypes: int64(3)
memory usage: 731.8 KB
```

```
In [18]: df.tweet_id.duplicated().sum()
```

```
Out[18]: 0
```

```
In [19]: df2.tweet_id.duplicated().sum()
```

```
Out[19]: 0
```

```
In [20]: df2.img_num.value_counts()
```

```
Out[20]: 1      1780
         2       198
         3        66
         4        31
         Name: img_num, dtype: int64
```

```
In [21]: 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]: #check duplicates ids in json tweets
sum(json_tweets.id.duplicated())

```

```

Out[22]: 28876

```

## 2.2 Quality

### 2.2.1 Df Dataframe

- Text column should not have links
- Timestamp column is in string
- missing name of the dogs
- dog's name are wrong
- The rating\_denominator should not be more than 10
- text column should not start with RT @dogs:
- "&" characters present in text

### 2.3 Df2 DataFrame

- The predicted breed of dogs should be in capital letter

## 3 tidiness

- json\_tweet column name id should be changed to tweet\_id to align it to other
- column name timestamp should be named tweet\_timestamp
- duplicate data on json\_tweet dataframe
- merger of all the dataframe

## 4 Clean

```

In [23]: #creating the duplicate file
df_clean=df.copy()
df2_clean=df2.copy()
json_tweets_clean=json_tweets.copy()

```

### 4.0.1 df: Text column should not have links

**Define: Link should be removed from text column**

#### 4.0.2 code

```
In [24]: #create a regex pattern in order to define the URLs that are present in the text column
link = r"http://S+"
```

```
In [25]: clean=[]
#loop for cleaning text column
for l in df['text']:
    clean.append(re.sub(link, '', l))
df_clean['text']=clean
```

#### 4.0.3 Test

```
In [26]: df['text'][1445]
```

```
Out[26]: 'Oh my god 10/10 for every little hot dog pupper'
```

#### 4.0.4 df: timestamp in string

**Define: timestamp from string to date**

#### 4.0.5 code

```
In [27]: date=[]
for d in df['timestamp']:
    date.append(datetime.strptime(d, "%Y-%m-%d %H:%M:%S +0000").date())
df_clean.timestamp=date
```

#### 4.0.6 Test

```
In [28]: df_clean.timestamp.sample(4)
```

```
Out[28]: 1862      2015-12-11
1202      2016-04-03
1199      2016-04-04
90        2017-06-13
Name: timestamp, dtype: object
```

#### 4.0.7 df: MISSING Dogs name

**Define: Remove the rows which has dogs name as None**

#### 4.0.8 code

```
In [29]: df_clean=df_clean[~(df_clean['name']=='None')]
```

#### 4.0.9 Test

```
In [30]: df_clean[df_clean['name']=='None']
```

```
Out[30]: Empty DataFrame
```

```
Columns: [tweet_id, in_reply_to_status_id, in_reply_to_user_id, timestamp, source, text]
Index: []
```

#### 4.0.10 df: Dogs name are wrong

**Define:** remove the rows whose dogs name are

#### 4.0.11 code

```
In [31]: name=[]
        for n in df.name:
            if(n.islower() and n not in name):
                name.append(n)
        print(name)
```

```
['such', 'a', 'quite', 'not', 'one', 'incredibly', 'mad', 'an', 'very', 'just', 'my', 'his', 'ac
```

```
In [32]: #create a df only with the first letter in names that is uppercase
        df_clean = df_clean[df_clean.name.map(lambda name: name.istitle())]
```

#### 4.0.12 Test

```
In [33]: df_clean[df_clean.name=='None']
```

```
Out[33]: Empty DataFrame
```

```
Columns: [tweet_id, in_reply_to_status_id, in_reply_to_user_id, timestamp, source, text]
Index: []
```

#### 4.0.13 Df: Rating denominator higher than 10

**Define:** Remove the lines that have a denominator higher than 10.

#### 4.0.14 Code

```
In [34]: df_clean['rating_denominator'].value_counts()
```

```
Out[34]: 10      1493
         50         1
         11         1
          7         1
         Name: rating_denominator, dtype: int64
```

```
In [35]: df_clean=df_clean[~(df_clean['rating_denominator']>10)]
```



## 4.1 Test

```
In [36]: df_clean.rating_denominator.value_counts()
```

```
Out[36]: 10    1493
         7      1
         Name: rating_denominator, dtype: int64
```

### 4.1.1 Df: Retweets present in the file: texts start with “RT @”

**Define:** Remove the text retweets starting with “RT @”, as we are looking only for original tweets.

#### 4.1.2 Code

```
In [37]: df_clean=df_clean[~(df_clean['text'].str.contains("RT @"))]
```

## 4.2 Test

```
In [38]: df_clean[df_clean.text.str.contains("RT @")]
```

```
Out[38]: Empty DataFrame
         Columns: [tweet_id, in_reply_to_status_id, in_reply_to_user_id, timestamp, source, text]
         Index: []
```

### 4.2.1 Df: “&” characters present in text

**Define:** Remove the characters “&” from text.

#### 4.2.2 Code

```
In [39]: df_clean=df_clean[~(df_clean['text'].str.contains("&"))]
```

## 4.3 Test

```
In [40]: df_clean[df_clean['text'].str.contains("&")]
```

```
Out[40]: Empty DataFrame
         Columns: [tweet_id, in_reply_to_status_id, in_reply_to_user_id, timestamp, source, text]
         Index: []
```

### 4.3.1 Df2: Some breed names have the first letter lowercase in p1, p2, p3 columns

**Define:** Align all the names. Make the first letter uppercase.

### 4.3.2 Code

```
In [41]: df2_clean.p1 = df2_clean.p1.str.title()
         df2_clean.p2 = df2_clean.p2.str.title()
         df2_clean.p3 = df2_clean.p3.str.title()
```

```
In [42]: df2_clean.head()
```

```
Out[42]:
```

	tweet_id	jpg_url	\
0	666020888022790149	https://pbs.twimg.com/media/CT4udnOWwAA0aMy.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	

	img_num	p1	p1_conf	p1_dog	p2	\
0	1	Welsh_Springer_Spaniel	0.465074	True	Collie	
1	1	Redbone	0.506826	True	Miniature_Pinscher	
2	1	German_Shepherd	0.596461	True	Malinois	
3	1	Rhodesian_Ridgeback	0.408143	True	Redbone	
4	1	Miniature_Pinscher	0.560311	True	Rottweiler	

	p2_conf	p2_dog	p3	p3_conf	p3_dog
0	0.156665	True	Shetland_Sheepdog	0.061428	True
1	0.074192	True	Rhodesian_Ridgeback	0.072010	True
2	0.138584	True	Bloodhound	0.116197	True
3	0.360687	True	Miniature_Pinscher	0.222752	True
4	0.243682	True	Doberman	0.154629	True

## 5 Tidiness

### 5.0.1 json\_tweet column name id should be changed to tweet\_id to align it to other

```
In [43]: json_tweets_clean=json_tweets_clean.rename(columns={'id':'tweet_id'})
```

```
In [44]: json_tweets_clean.head()
```

```
Out[44]:
```

	tweet_id	retweet_count	favorite_count
0	892420643555336193	8532	38588
1	892177421306343426	6270	33086
2	891815181378084864	4156	24912
3	891689557279858688	8660	42009
4	891327558926688256	9411	40140

### 5.0.2 In df DataFrame the column name timestamp should be changed to tweet\_timestamp

```
In [45]: df_clean=df_clean.rename(columns={'timestamp':'tweet_timestamp'})
```

```
In [46]: df_clean.head()
```

```

Out[46]:
      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

      tweet_timestamp                source  \
0  2017-08-01  <a href="http://twitter.com/download/iphone" r...
1  2017-08-01  <a href="http://twitter.com/download/iphone" r...
2  2017-07-31  <a href="http://twitter.com/download/iphone" r...
3  2017-07-30  <a href="http://twitter.com/download/iphone" r...
4  2017-07-29  <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

      retweeted_status_user_id  retweeted_status_timestamp  \
0                NaN                NaN
1                NaN                NaN
2                NaN                NaN
3                NaN                NaN
4                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

      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

```

### 5.0.3 eliminating duplicate data in json\_tweets dataframe

```

In [47]: #d=json_tweets_clean.duplicated()
         json_tweets_clean.drop_duplicates(subset='tweet_id',inplace=True)
         #json_tweets_clean=json_tweets_clean.loc[1:,d]

```

```
In [48]: json_tweets_clean.duplicated().sum()
```

```
Out[48]: 0
```

#### 5.0.4 merging of all 3 dataframe

```
In [49]: df_first=pd.merge(df_clean,df2_clean,on = 'tweet_id',how='left')
         df_master=pd.merge(df_first,json_tweets_clean,on = 'tweet_id',how='left')
```

```
In [50]: df_master.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1343 entries, 0 to 1342
Data columns (total 30 columns):
tweet_id                1343 non-null int64
in_reply_to_status_id   1 non-null float64
in_reply_to_user_id     1 non-null float64
tweet_timestamp         1343 non-null object
source                  1343 non-null object
text                    1343 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           1343 non-null object
rating_numerator        1343 non-null int64
rating_denominator      1343 non-null int64
name                    1343 non-null object
doggo                   1343 non-null object
floofer                 1343 non-null object
pupper                  1343 non-null object
puppo                   1343 non-null object
jpg_url                 1302 non-null object
img_num                 1302 non-null float64
p1                      1302 non-null object
p1_conf                 1302 non-null float64
p1_dog                  1302 non-null object
p2                      1302 non-null object
p2_conf                 1302 non-null float64
p2_dog                  1302 non-null object
p3                      1302 non-null object
p3_conf                 1302 non-null float64
p3_dog                  1302 non-null object
retweet_count           1342 non-null float64
favorite_count          1342 non-null float64
dtypes: float64(10), int64(3), object(17)
memory usage: 325.3+ KB
```

```
In [51]: #storing clean Data
         df_master.to_csv('twitter_archive_master.csv')
```

## 5.1 Analysis and Visualization

```
In [52]: #dataset
df_master
```

```
Out[52]:
```

	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	890971913173991426	NaN	NaN	
6	890609185150312448	NaN	NaN	
7	890240255349198849	NaN	NaN	
8	890006608113172480	NaN	NaN	
9	889880896479866881	NaN	NaN	
10	889638837579907072	NaN	NaN	
11	889531135344209921	NaN	NaN	
12	889278841981685760	NaN	NaN	
13	888917238123831296	NaN	NaN	
14	888804989199671297	NaN	NaN	
15	888554962724278272	NaN	NaN	
16	888078434458587136	NaN	NaN	
17	887705289381826560	NaN	NaN	
18	887473957103951883	NaN	NaN	
19	886983233522544640	NaN	NaN	
20	886736880519319552	NaN	NaN	
21	886680336477933568	NaN	NaN	
22	886366144734445568	NaN	NaN	
23	886258384151887873	NaN	NaN	
24	885984800019947520	NaN	NaN	
25	885528943205470208	NaN	NaN	
26	884925521741709313	NaN	NaN	
27	884876753390489601	NaN	NaN	
28	884562892145688576	NaN	NaN	
29	884162670584377345	NaN	NaN	
...	...	...	...	
1313	667534815156183040	NaN	NaN	
1314	667530908589760512	NaN	NaN	
1315	667502640335572993	NaN	NaN	
1316	667495797102141441	NaN	NaN	
1317	667455448082227200	NaN	NaN	
1318	667453023279554560	NaN	NaN	
1319	667405339315146752	NaN	NaN	
1320	667393430834667520	NaN	NaN	
1321	667211855547486208	NaN	NaN	
1322	667200525029539841	NaN	NaN	
1323	667182792070062081	NaN	NaN	
1324	667174963120574464	NaN	NaN	

1325	667171260800061440	NaN	NaN
1326	667165590075940865	NaN	NaN
1327	667160273090932737	NaN	NaN
1328	667119796878725120	NaN	NaN
1329	667090893657276420	NaN	NaN
1330	667073648344346624	NaN	NaN
1331	667062181243039745	NaN	NaN
1332	667012601033924608	NaN	NaN
1333	666996132027977728	NaN	NaN
1334	666817836334096384	NaN	NaN
1335	666804364988780544	NaN	NaN
1336	666776908487630848	NaN	NaN
1337	666739327293083650	NaN	NaN
1338	666691418707132416	NaN	NaN
1339	666644823164719104	NaN	NaN
1340	666454714377183233	NaN	NaN
1341	666447344410484738	NaN	NaN
1342	666418789513326592	NaN	NaN

	tweet_timestamp	source \
0	2017-08-01	<a href="http://twitter.com/download/iphone" r...
1	2017-08-01	<a href="http://twitter.com/download/iphone" r...
2	2017-07-31	<a href="http://twitter.com/download/iphone" r...
3	2017-07-30	<a href="http://twitter.com/download/iphone" r...
4	2017-07-29	<a href="http://twitter.com/download/iphone" r...
5	2017-07-28	<a href="http://twitter.com/download/iphone" r...
6	2017-07-27	<a href="http://twitter.com/download/iphone" r...
7	2017-07-26	<a href="http://twitter.com/download/iphone" r...
8	2017-07-26	<a href="http://twitter.com/download/iphone" r...
9	2017-07-25	<a href="http://twitter.com/download/iphone" r...
10	2017-07-25	<a href="http://twitter.com/download/iphone" r...
11	2017-07-24	<a href="http://twitter.com/download/iphone" r...
12	2017-07-24	<a href="http://twitter.com/download/iphone" r...
13	2017-07-23	<a href="http://twitter.com/download/iphone" r...
14	2017-07-22	<a href="http://twitter.com/download/iphone" r...
15	2017-07-22	<a href="http://twitter.com/download/iphone" r...
16	2017-07-20	<a href="http://twitter.com/download/iphone" r...
17	2017-07-19	<a href="http://twitter.com/download/iphone" r...
18	2017-07-19	<a href="http://twitter.com/download/iphone" r...
19	2017-07-17	<a href="http://twitter.com/download/iphone" r...
20	2017-07-16	<a href="http://twitter.com/download/iphone" r...
21	2017-07-16	<a href="http://twitter.com/download/iphone" r...
22	2017-07-15	<a href="http://twitter.com/download/iphone" r...
23	2017-07-15	<a href="http://twitter.com/download/iphone" r...
24	2017-07-14	<a href="http://twitter.com/download/iphone" r...
25	2017-07-13	<a href="http://twitter.com/download/iphone" r...
26	2017-07-12	<a href="http://twitter.com/download/iphone" r...
27	2017-07-11	<a href="http://twitter.com/download/iphone" r...

28	2017-07-11	<a href="http://twitter.com/download/iphone" r...
29	2017-07-09	<a href="http://twitter.com/download/iphone" r...
...	...	...
1313	2015-11-20	<a href="http://twitter.com" rel="nofollow">Tw...
1314	2015-11-20	<a href="http://twitter.com" rel="nofollow">Tw...
1315	2015-11-20	<a href="http://twitter.com" rel="nofollow">Tw...
1316	2015-11-20	<a href="http://twitter.com" rel="nofollow">Tw...
1317	2015-11-19	<a href="http://twitter.com" rel="nofollow">Tw...
1318	2015-11-19	<a href="http://twitter.com" rel="nofollow">Tw...
1319	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1320	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1321	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1322	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1323	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1324	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1325	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1326	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1327	2015-11-19	<a href="http://twitter.com/download/iphone" r...
1328	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1329	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1330	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1331	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1332	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1333	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1334	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1335	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1336	2015-11-18	<a href="http://twitter.com/download/iphone" r...
1337	2015-11-17	<a href="http://twitter.com/download/iphone" r...
1338	2015-11-17	<a href="http://twitter.com/download/iphone" r...
1339	2015-11-17	<a href="http://twitter.com/download/iphone" r...
1340	2015-11-17	<a href="http://twitter.com/download/iphone" r...
1341	2015-11-17	<a href="http://twitter.com/download/iphone" r...
1342	2015-11-17	<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	Meet Jax. He enjoys ice cream so much he gets ...	NaN
6	This is Zoey. She doesn't want to be one of th...	NaN
7	This is Cassie. She is a college pup. Studying...	NaN
8	This is Koda. He is a South Australian decksha...	NaN
9	This is Bruno. He is a service shark. Only get...	NaN
10	This is Ted. He does his best. Sometimes that'...	NaN
11	This is Stuart. He's sporting his favorite fan...	NaN
12	This is Oliver. You're witnessing one of his m...	NaN

13	This is Jim. He found a fren. Taught him how t...	NaN
14	This is Zeke. He has a new stick. Very proud o...	NaN
15	This is Ralphus. He's powering up. Attempting ...	NaN
16	This is Gerald. He was just told he didn't get...	NaN
17	This is Jeffrey. He has a monopoly on the pool...	NaN
18	This is Canela. She attempted some fancy porch...	NaN
19	This is Maya. She's very shy. Rarely leaves he...	NaN
20	This is Mingus. He's a wonderful father to his...	NaN
21	This is Derek. He's late for a dog meeting. 13...	NaN
22	This is Roscoe. Another pupper fallen victim t...	NaN
23	This is Waffles. His doggles are pupside down...	NaN
24	Viewer discretion advised. This is Jimbo. He w...	NaN
25	This is Maisey. She fell asleep mid-excavation...	NaN
26	This is Earl. He found a hat. Nervous about wh...	NaN
27	This is Lola. It's her first time outside. Mus...	NaN
28	This is Kevin. He's just so happy. 13/10 what ...	NaN
29	Meet Yogi. He doesn't have any important dog m...	NaN
...	...	...
1313	This is Frank (pronounced "Fronq"). Too many b...	NaN
1314	Meet Naphaniel. He doesn't necessarily enjoy h...	NaN
1315	Say hello to Hall and Oates. Oates is winking ...	NaN
1316	This is Philippe from Soviet Russia. Commandin...	NaN
1317	This is Reese and Twips. Reese protects Twips...	NaN
1318	Meet Cupcake. I would do unspeakable things fo...	NaN
1319	This is Biden. Biden just tripped... 7/10 http...	NaN
1320	This is Fwed. He is a Canadian Asian Taylormad...	NaN
1321	This is Genevieve. She is a golden retriever c...	NaN
1322	This is Joshwa. He is a fuckboy supreme. He cl...	NaN
1323	This is Timison. He just told an awful joke bu...	NaN
1324	This is Clarence. His face says he doesn't wan...	NaN
1325	Say hello to Kenneth. He likes Reese's Puffs. ...	NaN
1326	This is Churlie. AKA Fetty Woof. Lost eye savi...	NaN
1327	This is Bradlay. He is a Ronaldinho Matsuyama ...	NaN
1328	This is Gabe. He is a southern Baklava. Gabe h...	NaN
1329	This is Clybe. He is an Anemone Valdez. One ea...	NaN
1330	Here is Dave. He is actually just a skinny leg...	NaN
1331	This is Keet. He is a Floridian Amukamara. Abs...	NaN
1332	This is Klevin. He laughs a lot. Very cool dog...	NaN
1333	This is Carll. He wants to be a donkey. But al...	NaN
1334	This is Jeph. He is a German Boston Shuttlecoc...	NaN
1335	This is Jockson. He is a Pinnacle Sagittarius...	NaN
1336	This is Josep. He is a Rye Manganese mix. Can ...	NaN
1337	This is Lugan. He is a Bohemian Rhapsody. Very...	NaN
1338	This is Christoper. He is a spotted Penne. Can...	NaN
1339	This is Jimothy. He is a Botwanian Gouda. Can ...	NaN
1340	I'll name the dogs from now on. This is Kreggo...	NaN
1341	This is Scout. She is a black Downton Abbey. I...	NaN
1342	This is Walter. He is an Alaskan Terrapin. Lov...	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	NaN	NaN
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
...	...	...
1313	NaN	NaN
1314	NaN	NaN
1315	NaN	NaN
1316	NaN	NaN
1317	NaN	NaN
1318	NaN	NaN
1319	NaN	NaN
1320	NaN	NaN
1321	NaN	NaN
1322	NaN	NaN
1323	NaN	NaN
1324	NaN	NaN
1325	NaN	NaN
1326	NaN	NaN
1327	NaN	NaN

1328	NaN	NaN
1329	NaN	NaN
1330	NaN	NaN
1331	NaN	NaN
1332	NaN	NaN
1333	NaN	NaN
1334	NaN	NaN
1335	NaN	NaN
1336	NaN	NaN
1337	NaN	NaN
1338	NaN	NaN
1339	NaN	NaN
1340	NaN	NaN
1341	NaN	NaN
1342	NaN	NaN

	expanded_urls	...	\
0	https://twitter.com/dog_rates/status/892420643...	...	
1	https://twitter.com/dog_rates/status/892177421...	...	
2	https://twitter.com/dog_rates/status/891815181...	...	
3	https://twitter.com/dog_rates/status/891689557...	...	
4	https://twitter.com/dog_rates/status/891327558...	...	
5	https://gofundme.com/ydvmve-surgery-for-jax,ht...	...	
6	https://twitter.com/dog_rates/status/890609185...	...	
7	https://twitter.com/dog_rates/status/890240255...	...	
8	https://twitter.com/dog_rates/status/890006608...	...	
9	https://twitter.com/dog_rates/status/889880896...	...	
10	https://twitter.com/dog_rates/status/889638837...	...	
11	https://twitter.com/dog_rates/status/889531135...	...	
12	https://twitter.com/dog_rates/status/889278841...	...	
13	https://twitter.com/dog_rates/status/888917238...	...	
14	https://twitter.com/dog_rates/status/888804989...	...	
15	https://twitter.com/dog_rates/status/888554962...	...	
16	https://twitter.com/dog_rates/status/888078434...	...	
17	https://twitter.com/dog_rates/status/887705289...	...	
18	https://twitter.com/dog_rates/status/887473957...	...	
19	https://twitter.com/dog_rates/status/886983233...	...	
20	https://www.gofundme.com/mingusneedsus,https:/...	...	
21	https://twitter.com/dog_rates/status/886680336...	...	
22	https://twitter.com/dog_rates/status/886366144...	...	
23	https://twitter.com/dog_rates/status/886258384...	...	
24	https://twitter.com/dog_rates/status/885984800...	...	
25	https://twitter.com/dog_rates/status/885528943...	...	
26	https://twitter.com/dog_rates/status/884925521...	...	
27	https://twitter.com/dog_rates/status/884876753...	...	
28	https://twitter.com/dog_rates/status/884562892...	...	
29	https://twitter.com/dog_rates/status/884162670...	...	
...	...	...	

1313 https://twitter.com/dog\_rates/status/667534815...  
 1314 https://twitter.com/dog\_rates/status/667530908...  
 1315 https://twitter.com/dog\_rates/status/667502640...  
 1316 https://twitter.com/dog\_rates/status/667495797...  
 1317 https://twitter.com/dog\_rates/status/667455448...  
 1318 https://twitter.com/dog\_rates/status/667453023...  
 1319 https://twitter.com/dog\_rates/status/667405339...  
 1320 https://twitter.com/dog\_rates/status/667393430...  
 1321 https://twitter.com/dog\_rates/status/667211855...  
 1322 https://twitter.com/dog\_rates/status/667200525...  
 1323 https://twitter.com/dog\_rates/status/667182792...  
 1324 https://twitter.com/dog\_rates/status/667174963...  
 1325 https://twitter.com/dog\_rates/status/667171260...  
 1326 https://twitter.com/dog\_rates/status/667165590...  
 1327 https://twitter.com/dog\_rates/status/667160273...  
 1328 https://twitter.com/dog\_rates/status/667119796...  
 1329 https://twitter.com/dog\_rates/status/667090893...  
 1330 https://twitter.com/dog\_rates/status/667073648...  
 1331 https://twitter.com/dog\_rates/status/667062181...  
 1332 https://twitter.com/dog\_rates/status/667012601...  
 1333 https://twitter.com/dog\_rates/status/666996132...  
 1334 https://twitter.com/dog\_rates/status/666817836...  
 1335 https://twitter.com/dog\_rates/status/666804364...  
 1336 https://twitter.com/dog\_rates/status/666776908...  
 1337 https://twitter.com/dog\_rates/status/666739327...  
 1338 https://twitter.com/dog\_rates/status/666691418...  
 1339 https://twitter.com/dog\_rates/status/666644823...  
 1340 https://twitter.com/dog\_rates/status/666454714...  
 1341 https://twitter.com/dog\_rates/status/666447344...  
 1342 https://twitter.com/dog\_rates/status/666418789...

	p1_conf	p1_dog	p2	p2_conf	p2_dog \
0	0.097049	False	Bagel	0.085851	False
1	0.323581	True	Pekinese	0.090647	True
2	0.716012	True	Malamute	0.078253	True
3	0.170278	False	Labrador_Retriever	0.168086	True
4	0.555712	True	English_Springer	0.225770	True
5	0.341703	True	Border_Collie	0.199287	True
6	0.487574	True	Irish_Setter	0.193054	True
7	0.511319	True	Cardigan	0.451038	True
8	0.957979	True	Pomeranian	0.013884	True
9	0.377417	True	Labrador_Retriever	0.151317	True
10	0.991650	True	Boxer	0.002129	True
11	0.953442	True	Labrador_Retriever	0.013834	True
12	0.626152	True	Borzoi	0.194742	True
13	0.714719	True	Tibetan_Mastiff	0.120184	True
14	0.469760	True	Labrador_Retriever	0.184172	True
15	0.700377	True	Eskimo_Dog	0.166511	True

16	0.995026	True	Pug	0.000932	True
17	0.821664	True	Redbone	0.087582	True
18	0.809197	True	Rhodesian_Ridgeback	0.054950	True
19	0.793469	True	Toy_Terrier	0.143528	True
20	0.309706	True	Great_Pyrenees	0.186136	True
21	0.738995	False	Sports_Car	0.139952	False
22	0.999201	True	Chihuahua	0.000361	True
23	0.943575	True	Shower_Cap	0.025286	False
24	0.972494	True	Shih-Tzu	0.006630	True
25	0.369275	True	Labrador_Retriever	0.265835	True
26	0.259916	True	American_Staffordshire_Terrier	0.198451	True
27	0.822103	True	Norwich_Terrier	0.106075	True
28	0.546406	True	French_Bulldog	0.404291	True
29	0.707046	True	Malinois	0.199396	True
...	...	...	...	...	...
1313	0.435254	True	Cardigan	0.307407	True
1314	0.633037	True	Kuvasz	0.146391	True
1315	0.996709	True	Golden_Retriever	0.001688	True
1316	0.143957	True	Christmas_Stocking	0.118651	False
1317	0.676376	True	Irish_Terrier	0.054933	True
1318	0.825670	True	French_Bulldog	0.056639	True
1319	0.381377	True	Leonberg	0.127998	True
1320	0.557009	True	Border_Collie	0.271963	True
1321	0.462556	True	Labrador_Retriever	0.454937	True
1322	0.694904	True	Malamute	0.232006	True
1323	0.949892	True	Irish_Setter	0.010564	True
1324	0.266437	True	Chihuahua	0.243223	True
1325	0.841265	True	Lakeland_Terrier	0.052744	True
1326	0.140173	True	Rottweiler	0.134094	True
1327	0.471351	True	Miniature_Poodle	0.091992	True
1328	0.741563	True	Chihuahua	0.057866	True
1329	0.959514	True	Italian_Greyhound	0.005370	True
1330	0.483682	True	Pug	0.092494	True
1331	0.825678	True	Vizsla	0.090998	True
1332	0.987230	False	African_Hunting_Dog	0.012601	False
1333	0.507637	False	Rottweiler	0.062490	True
1334	0.496953	True	Standard_Schnauzer	0.285276	True
1335	0.328792	True	Brittany_Spaniel	0.283545	True
1336	0.375057	False	Miniature_Pinscher	0.167175	True
1337	0.546933	True	Cocker_Spaniel	0.165255	True
1338	0.975401	True	Beagle	0.008687	True
1339	0.044333	True	Pembroke	0.043209	True
1340	0.278954	True	Labrador_Retriever	0.237612	True
1341	0.322084	True	Giant_Schnauzer	0.287955	True
1342	0.149680	True	Papillon	0.148258	True

	p3	p3_conf	p3_dog	retweet_count	\
0	Banana	0.076110	False	8532.0	

1	Papillon	0.068957	True	6270.0
2	Kelpie	0.031379	True	4156.0
3	Spatula	0.040836	False	8660.0
4	German_Short-Haired_Pointer	0.175219	True	9411.0
5	Ice_Lolly	0.193548	False	2073.0
6	Chesapeake_Bay_Retriever	0.118184	True	4270.0
7	Chihuahua	0.029248	True	7427.0
8	Chow	0.008167	True	7341.0
9	Muzzle	0.082981	False	4976.0
10	Staffordshire_Bullterrier	0.001498	True	4548.0
11	Redbone	0.007958	True	2236.0
12	Saluki	0.027351	True	5419.0
13	Labrador_Retriever	0.105506	True	4501.0
14	English_Setter	0.073482	True	4339.0
15	Malamute	0.111411	True	3580.0
16	Bull_Mastiff	0.000903	True	3485.0
17	Weimaraner	0.026236	True	5389.0
18	Beagle	0.038915	True	18199.0
19	Can_Opener	0.032253	False	7768.0
20	Dandie_Dinmont	0.086346	True	3291.0
21	Car_Wheel	0.044173	False	4457.0
22	Boston_Bull	0.000076	True	3196.0
23	Siamese_Cat	0.002849	False	6278.0
24	Bernese_Mountain_Dog	0.006239	True	6810.0
25	Kuvasz	0.134697	True	6424.0
26	Staffordshire_Bullterrier	0.127725	True	18426.0
27	Norfolk_Terrier	0.037348	True	5788.0
28	Brabancon_Griffon	0.044002	True	4865.0
29	Norwegian_Elkhound	0.049148	True	2982.0
...	...	...	...	...
1313	Cocker_Spaniel	0.033158	True	556.0
1314	Labrador_Retriever	0.046184	True	251.0
1315	Beagle	0.000712	True	224.0
1316	Ski_Mask	0.092482	False	283.0
1317	Yorkshire_Terrier	0.040576	True	63.0
1318	Staffordshire_Bullterrier	0.054018	True	92.0
1319	Golden_Retriever	0.069357	True	224.0
1320	Collie	0.073473	True	58.0
1321	Kuvasz	0.024193	True	248.0
1322	Eskimo_Dog	0.050635	True	269.0
1323	Chesapeake_Bay_Retriever	0.005821	True	6412.0
1324	Bluetick	0.072806	True	85.0
1325	Irish_Water_Spaniel	0.034402	True	92.0
1326	Beagle	0.081900	True	1180.0
1327	Standard_Poodle	0.087385	True	64.0
1328	Toy_Poodle	0.039125	True	131.0
1329	Pomeranian	0.002641	True	129.0
1330	Brabancon_Griffon	0.057495	True	125.0

1331	Kelpie	0.022956	True	55.0
1332	Coyote	0.000057	False	227.0
1333	Water_Buffalo	0.048425	False	96.0
1334	Giant_Schnauzer	0.073764	True	258.0
1335	Ibizan_Hound	0.057461	True	94.0
1336	Chihuahua	0.086951	True	176.0
1337	Toy_Poodle	0.095959	True	66.0
1338	Bloodhound	0.005394	True	47.0
1339	West_Highland_White_Terrier	0.038906	True	83.0
1340	Great_Pyrenees	0.171106	True	210.0
1341	Labrador_Retriever	0.166331	True	20.0
1342	Chihuahua	0.142860	True	46.0

	favorite_count
0	38588.0
1	33086.0
2	24912.0
3	42009.0
4	40140.0
5	11797.0
6	27667.0
7	31787.0
8	30543.0
9	27651.0
10	27046.0
11	15030.0
12	25167.0
13	28948.0
14	25460.0
15	19809.0
16	21665.0
17	30045.0
18	68794.0
19	34984.0
20	12019.0
21	22316.0
22	21096.0
23	27851.0
24	32550.0
25	35828.0
26	77188.0
27	27807.0
28	24189.0
29	20291.0
...	...
1313	839.0
1314	485.0
1315	543.0

1316	538.0
1317	194.0
1318	320.0
1319	474.0
1320	198.0
1321	500.0
1322	626.0
1323	14647.0
1324	251.0
1325	225.0
1326	2741.0
1327	262.0
1328	334.0
1329	340.0
1330	411.0
1331	221.0
1332	456.0
1333	248.0
1334	527.0
1335	243.0
1336	356.0
1337	239.0
1338	190.0
1339	234.0
1340	523.0
1341	105.0
1342	126.0

[1343 rows x 30 columns]

In [53]: `df_master.duplicated().sum()`

Out[53]: 0

In [54]: *#columns that have missing values*  
`df_master.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1343 entries, 0 to 1342
Data columns (total 30 columns):
tweet_id                1343 non-null int64
in_reply_to_status_id    1 non-null float64
in_reply_to_user_id      1 non-null float64
tweet_timestamp          1343 non-null object
source                   1343 non-null object
text                     1343 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          1343 non-null object
rating_numerator       1343 non-null int64
rating_denominator     1343 non-null int64
name                   1343 non-null object
doggo                  1343 non-null object
floofer                1343 non-null object
pupper                1343 non-null object
puppo                  1343 non-null object
jpg_url                1302 non-null object
img_num                1302 non-null float64
p1                     1302 non-null object
p1_conf                1302 non-null float64
p1_dog                 1302 non-null object
p2                     1302 non-null object
p2_conf                1302 non-null float64
p2_dog                 1302 non-null object
p3                     1302 non-null object
p3_conf                1302 non-null float64
p3_dog                 1302 non-null object
retweet_count          1342 non-null float64
favorite_count          1342 non-null float64
dtypes: float64(10), int64(3), object(17)
memory usage: 325.3+ KB

```

```

In [55]: c=['in_reply_to_status_id','in_reply_to_user_id','retweeted_status_id','retweeted_status_id']
df_master.drop(['in_reply_to_status_id','in_reply_to_user_id','retweeted_status_id','retweeted_status_id'],inplace=True)

```

```

In [56]: df_master.info()

```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1343 entries, 0 to 1342
Data columns (total 25 columns):
tweet_id          1343 non-null int64
tweet_timestamp   1343 non-null object
source            1343 non-null object
text              1343 non-null object
expanded_urls     1343 non-null object
rating_numerator  1343 non-null int64
rating_denominator 1343 non-null int64
name              1343 non-null object
doggo             1343 non-null object
floofer           1343 non-null object
pupper            1343 non-null object
puppo             1343 non-null object
jpg_url           1302 non-null object
img_num           1302 non-null float64
p1                1302 non-null object

```



```

p1_conf          1302 non-null float64
p1_dog           1302 non-null object
p2               1302 non-null object
p2_conf          1302 non-null float64
p2_dog           1302 non-null object
p3               1302 non-null object
p3_conf          1302 non-null float64
p3_dog           1302 non-null object
retweet_count    1342 non-null float64
favorite_count   1342 non-null float64
dtypes: float64(6), int64(3), object(16)
memory usage: 272.8+ KB

```

### 5.1.1 Descriptive Statistics

```
In [57]: df_master.describe()
```

```

Out[57]:

```

	tweet_id	rating_numerator	rating_denominator	img_num	\
count	1.343000e+03	1343.000000	1343.0	1302.000000	
mean	7.454961e+17	12.121370	10.0	1.211982	
std	6.837014e+16	48.239137	0.0	0.575902	
min	6.664188e+17	2.000000	10.0	1.000000	
25%	6.812179e+17	10.000000	10.0	1.000000	
50%	7.276857e+17	11.000000	10.0	1.000000	
75%	8.011415e+17	12.000000	10.0	1.000000	
max	8.924206e+17	1776.000000	10.0	4.000000	

	p1_conf	p2_conf	p3_conf	retweet_count	favorite_count
count	1302.000000	1.302000e+03	1.302000e+03	1342.000000	1342.000000
mean	0.602612	1.341769e-01	6.017569e-02	2717.324143	9350.049180
std	0.272807	1.009687e-01	5.134916e-02	3980.101846	11885.148102
min	0.044333	1.011300e-08	1.740170e-10	20.000000	105.000000
25%	0.368586	5.206585e-02	1.532280e-02	710.750000	2341.000000
50%	0.601799	1.191595e-01	4.901590e-02	1547.500000	4850.500000
75%	0.867487	1.978880e-01	9.394012e-02	3324.250000	12108.250000
max	1.000000	4.880140e-01	2.710420e-01	60674.000000	123762.000000

```

In [58]: #spices of dogs
df_master.p1.value_counts()

```

```

Out[58]: Golden_Retriever      88
Pembroke                      64
Labrador_Retriever            60
Chihuahua                     58
Pug                           38
Chow                          31
Toy_Poodle                    26
Pomeranian                    24

```

Samoyed	23
Malamute	21
French_Bulldog	21
Cocker_Spaniel	18
Siberian_Husky	18
Miniature_Pinscher	16
Chesapeake_Bay_Retriever	16
Staffordshire_Bullterrier	15
Shih-Tzu	15
Cardigan	14
Seat_Belt	14
German_Shepherd	14
Beagle	13
Shetland_Sheepdog	13
Eskimo_Dog	12
Pekinese	11
Teddy	11
Border_Collie	10
American_Staffordshire_Terrier	10
Vizsla	10
Rottweiler	9
Airedale	9
..	..
Lawn_Mower	1
Cougar	1
Tailed_Frog	1
Crane	1
Banana	1
Bustard	1
Boathouse	1
Guenon	1
Hand_Blower	1
Maillot	1
Cuirass	1
Pitcher	1
Otter	1
Slug	1
Timber_Wolf	1
Seashore	1
American_Black_Bear	1
Tricycle	1
Common_Iguana	1
Nail	1
Wild_Boar	1
African_Grey	1
Wooden_Spoon	1
Agama	1
Cash_Machine	1

Leopard	1
Damselfly	1
Leatherback_Turtle	1
Vacuum	1
Pedestal	1

Name: p1, Length: 286, dtype: int64

```
In [59]: #spices of dogs
df_master.p2.value_counts()
```

```
Out[59]: Labrador_Retriever      63
Golden_Retriever      59
Cardigan      54
French_Bulldog      28
Toy_Poodle      26
Chihuahua      26
Pomeranian      25
Chesapeake_Bay_Retriever      24
Miniature_Poodle      23
Pembroke      23
Eskimo_Dog      21
Cocker_Spaniel      21
Siberian_Husky      19
Collie      19
Beagle      17
Boston_Bull      17
Norwegian_Elkhound      16
Miniature_Pinscher      15
Staffordshire_Bullterrier      15
Pekinese      14
American_Staffordshire_Terrier      14
Italian_Greyhound      14
Kelpie      13
Kuvasz      12
Samoyed      12
Pug      12
Basset      11
Toy_Terrier      11
Malinois      11
Irish_Terrier      11
..
Lawn_Mower      1
Crutch      1
Indian_Elephant      1
Hen      1
China_Cabinet      1
Water_Bottle      1
Tick      1
```

Bernese_Mountain_Dog	1
Ox	1
Orange	1
Academic_Gown	1
Umbrella	1
Shoji	1
Space_Heater	1
Tow_Truck	1
Banded_Gecko	1
Sunglass	1
Oscilloscope	1
Toucan	1
Dining_Table	1
Sleeping_Bag	1
Maillot	1
Bow_Tie	1
Barrel	1
Kerry_Blue_Terrier	1
Snail	1
Giant_Panda	1
Cash_Machine	1
Television	1
Rhinoceros_Beetle	1

Name: p2, Length: 298, dtype: int64

```
In [60]: #spices of dogs
df_master.p3.value_counts()
```

Labrador_Retriever	54
Chihuahua	38
Kelpie	28
Golden_Retriever	26
Chow	25
Beagle	22
Malamute	22
Toy_Poodle	22
Pembroke	22
Eskimo_Dog	21
Pekinese	21
Chesapeake_Bay_Retriever	20
French_Bulldog	20
Staffordshire_Bullterrier	20
Pomeranian	18
Cocker_Spaniel	18
Kuvasz	18
Toy_Terrier	16
Pug	15
American_Staffordshire_Terrier	15

Bull_Mastiff	15
Siberian_Husky	14
Cardigan	14
Lakeland_Terrier	14
Great_Pyrenees	14
Basenji	13
Norfolk_Terrier	12
Doormat	11
Siamese_Cat	11
Boston_Bull	11
..	
Joystick	1
Meerkat	1
Shoji	1
Lakeside	1
Standard_Schnauzer	1
Swimming_Trunks	1
Wig	1
Cheetah	1
Hamster	1
Nail	1
Sliding_Door	1
Bullfrog	1
Whiptail	1
Mask	1
Bow	1
Wallaby	1
Sussex_Spaniel	1
Traffic_Light	1
Mink	1
Abaya	1
Plastic_Bag	1
Green_Lizard	1
Bagel	1
Pillow	1
Otter	1
Neck_Brace	1
Chest	1
Sealyham_Terrier	1
Red_Wolf	1
Parallel_Bars	1

Name: p3, Length: 301, dtype: int64

```
In [61]: #count the number of ratings numerator
df_master['rating_numerator'].value_counts()
```

```
Out[61]: 12      333
        11      299
```

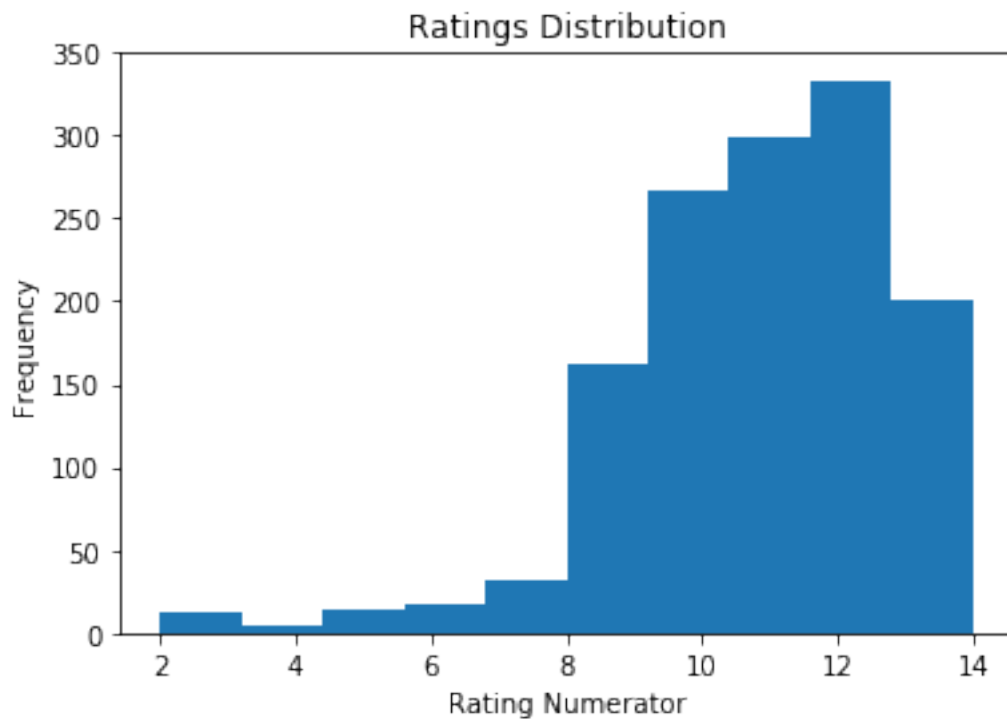
10	266
13	183
9	93
8	69
7	32
14	17
6	17
5	14
3	10
4	5
2	2
1776	1
75	1
27	1

Name: rating\_numerator, dtype: int64

**The highest rating is 75/10 while lowest is 2/10 while there are also outliers 1776/10**

```
In [62]: #plot the count of rating numerators between 2 and 75
plt.hist(df_master['rating_numerator'], range = (2, 14));
plt.xlabel('Rating Numerator')
plt.ylabel('Frequency')
plt.title('Ratings Distribution')
```

```
Out[62]: Text(0.5,1,'Ratings Distribution')
```



### 5.1.2 The graph clearly shows that it is left skewed with peak at 12

In [63]: *#check outlinears*

```
df_master[df_master.rating_numerator == 1776]
```

```
Out[63]:
```

	tweet_id	tweet_timestamp		source		text		expanded_urls	rating_numerator	
577	749981277374128128	2016-07-04		<a href="https://about.twitter.com/products/tw...		This is Atticus. He's quite simply America af...		https://twitter.com/dog_rates/status/749981277...	1776	
	rating_denominator	name	doggo	floofer	...	p1_conf				
577	10	Atticus	None	None	...	0.533941				
	p1_dog	p2	p2_conf	p2_dog	p3	p3_conf	p3_dog			
577	False	Sunglasses	0.080822	False	Sunglass	0.050776	False			
	retweet_count	favorite_count								
577	2693.0	5482.0								

[1 rows x 25 columns]

In [64]: *#check dogs with rating 2*

```
df_master[df_master.rating_numerator == 2]
```

```
Out[64]:
```

	tweet_id	tweet_timestamp		source		text		expanded_urls	rating_numerator	
1049	678424312106393600	2015-12-20		<a href="http://twitter.com/download/iphone" r...		This is Crystal. She's a shitty fireman. No se...		https://twitter.com/dog_rates/status/678424312...	2	
1110	674265582246694913	2015-12-08		<a href="http://twitter.com/download/iphone" r...		This is Henry. He's a shit dog. Short pointy e...		https://twitter.com/dog_rates/status/674265582...	2	
	rating_denominator	name	doggo	floofer	...	p1_conf				
1049	10	Crystal	None	None	...	0.759945				

1110		10	Henry	None	None	...	0.998075
------	--	----	-------	------	------	-----	----------

	p1_dog		p2	p2_conf	p2_dog		p3	p3_conf	p3_dog	\
1049	True	Toy_Poodle	0.101194	True	Shih-Tzu	0.056037	True			
1110	False	Ice_Lolly	0.000984	False	Leafhopper	0.000097	False			

	retweet_count	favorite_count
1049	2770.0	5741.0
1110	896.0	1699.0

[2 rows x 25 columns]

In [65]: *#check dogs with rating 75*

```
df_master[df_master.rating_numerator == 75]
```

Out [65]:

	tweet_id	tweet_timestamp	\
402	786709082849828864	2016-10-13	
		source	\
402	<a href="http://twitter.com/download/iphone" r...		
		text	\
402	This is Logan, the Chow who lived. He solemnly...		
		expanded_urls	rating_numerator \
402	https://twitter.com/dog_rates/status/786709082...		75
	rating_denominator	name	doggo floofer ... p1_conf p1_dog \
402	10	Logan	None None ... 0.467321 True
	p2	p2_conf	p2_dog p3 p3_conf p3_dog retweet_count \
402	Persian_Cat	0.122978	False Chow 0.102654 True 6810.0
	favorite_count		
402	19789.0		

[1 rows x 25 columns]

In [66]: *#check dogs with rating 12*

```
df_master[df_master.rating_numerator == 12]
```

Out [66]:

	tweet_id	tweet_timestamp	\
2	891815181378084864	2017-07-31	
4	891327558926688256	2017-07-29	
10	889638837579907072	2017-07-25	
13	888917238123831296	2017-07-23	
16	888078434458587136	2017-07-20	
22	886366144734445568	2017-07-15	
24	885984800019947520	2017-07-14	



26	884925521741709313	2017-07-12
29	884162670584377345	2017-07-09
30	883838122936631299	2017-07-09
34	882762694511734784	2017-07-06
38	881906580714921986	2017-07-03
40	881268444196462592	2017-07-01
42	880465832366813184	2017-06-29
43	880221127280381952	2017-06-29
45	879492040517615616	2017-06-27
47	879376492567855104	2017-06-26
54	877556246731214848	2017-06-21
56	877201837425926144	2017-06-20
57	876838120628539392	2017-06-19
62	874680097055178752	2017-06-13
65	873213775632977920	2017-06-09
68	872122724285648897	2017-06-06
70	871762521631449091	2017-06-05
71	871515927908634625	2017-06-04
79	869596645499047938	2017-05-30
82	868552278524837888	2017-05-27
84	867421006826221569	2017-05-24
85	866686824827068416	2017-05-22
87	866334964761202691	2017-05-21
...	...	...
972	684147889187209216	2016-01-04
995	681981167097122816	2015-12-29
998	681654059175129088	2015-12-29
1000	681523177663676416	2015-12-28
1013	680836378243002368	2015-12-26
1017	680497766108381184	2015-12-25
1019	680206703334408192	2015-12-25
1020	680161097740095489	2015-12-24
1055	678021115718029313	2015-12-19
1082	675710890956750848	2015-12-12
1086	675362609739206656	2015-12-11
1094	675003128568291329	2015-12-10
1103	674422304705744896	2015-12-09
1117	674038233588723717	2015-12-08
1121	673708611235921920	2015-12-07
1128	673662677122719744	2015-12-07
1142	672997845381865473	2015-12-05
1161	672481316919734272	2015-12-03
1163	672466075045466113	2015-12-03
1167	672245253877968896	2015-12-03
1208	670823764196741120	2015-11-29
1230	670421925039075328	2015-11-28
1232	670411370698022913	2015-11-28
1236	670374371102445568	2015-11-27

1254	669625907762618368	2015-11-25
1267	669328503091937280	2015-11-25
1272	669006782128353280	2015-11-24
1277	668979806671884288	2015-11-24
1292	668528771708952576	2015-11-22
1300	668190681446379520	2015-11-21

```

source \
2    <a href="http://twitter.com/download/iphone" r...
4    <a href="http://twitter.com/download/iphone" r...
10   <a href="http://twitter.com/download/iphone" r...
13   <a href="http://twitter.com/download/iphone" r...
16   <a href="http://twitter.com/download/iphone" r...
22   <a href="http://twitter.com/download/iphone" r...
24   <a href="http://twitter.com/download/iphone" r...
26   <a href="http://twitter.com/download/iphone" r...
29   <a href="http://twitter.com/download/iphone" r...
30   <a href="http://twitter.com/download/iphone" r...
34   <a href="http://twitter.com/download/iphone" r...
38   <a href="http://twitter.com/download/iphone" r...
40   <a href="http://twitter.com/download/iphone" r...
42   <a href="http://twitter.com/download/iphone" r...
43   <a href="http://twitter.com/download/iphone" r...
45   <a href="http://twitter.com/download/iphone" r...
47   <a href="http://twitter.com/download/iphone" r...
54   <a href="http://twitter.com/download/iphone" r...
56   <a href="http://twitter.com/download/iphone" r...
57   <a href="http://twitter.com/download/iphone" r...
62   <a href="http://twitter.com/download/iphone" r...
65   <a href="http://twitter.com/download/iphone" r...
68   <a href="http://twitter.com/download/iphone" r...
70   <a href="http://twitter.com/download/iphone" r...
71   <a href="http://twitter.com/download/iphone" r...
79   <a href="http://twitter.com/download/iphone" r...
82   <a href="http://twitter.com/download/iphone" r...
84   <a href="http://twitter.com/download/iphone" r...
85   <a href="http://twitter.com/download/iphone" r...
87   <a href="http://twitter.com/download/iphone" r...
...
972  <a href="http://vine.co" rel="nofollow">Vine -...
995  <a href="http://twitter.com/download/iphone" r...
998  <a href="http://twitter.com/download/iphone" r...
1000 <a href="http://twitter.com/download/iphone" r...
1013 <a href="http://twitter.com/download/iphone" r...
1017 <a href="http://twitter.com/download/iphone" r...
1019 <a href="http://twitter.com/download/iphone" r...
1020 <a href="http://twitter.com/download/iphone" r...
1055 <a href="http://twitter.com/download/iphone" r...

```

1082 <a href="http://twitter.com/download/iphone" r...  
 1086 <a href="http://twitter.com/download/iphone" r...  
 1094 <a href="http://twitter.com/download/iphone" r...  
 1103 <a href="http://twitter.com/download/iphone" r...  
 1117 <a href="http://twitter.com/download/iphone" r...  
 1121 <a href="http://twitter.com/download/iphone" r...  
 1128 <a href="http://twitter.com/download/iphone" r...  
 1142 <a href="http://twitter.com/download/iphone" r...  
 1161 <a href="http://twitter.com/download/iphone" r...  
 1163 <a href="http://twitter.com/download/iphone" r...  
 1167 <a href="http://twitter.com/download/iphone" r...  
 1208 <a href="http://twitter.com/download/iphone" r...  
 1230 <a href="http://twitter.com/download/iphone" r...  
 1232 <a href="http://twitter.com/download/iphone" r...  
 1236 <a href="http://twitter.com/download/iphone" r...  
 1254 <a href="http://twitter.com/download/iphone" r...  
 1267 <a href="http://twitter.com/download/iphone" r...  
 1272 <a href="http://twitter.com/download/iphone" r...  
 1277 <a href="http://twitter.com/download/iphone" r...  
 1292 <a href="http://twitter.com/download/iphone" r...  
 1300 <a href="http://twitter.com/download/iphone" r...

text \

2 This is Archie. He is a rare Norwegian Pouncin...  
 4 This is Franklin. He would like you to stop ca...  
 10 This is Ted. He does his best. Sometimes that'...  
 13 This is Jim. He found a fren. Taught him how t...  
 16 This is Gerald. He was just told he didn't get...  
 22 This is Roscoe. Another pupper fallen victim t...  
 24 Viewer discretion advised. This is Jimbo. He w...  
 26 This is Earl. He found a hat. Nervous about wh...  
 29 Meet Yogi. He doesn't have any important dog m...  
 30 This is Noah. He can't believe someone made th...  
 34 This is Gus. He's quite the cheeky pupper. Alr...  
 38 This is Rey. He's a Benebop Cumberfloof. 12/10...  
 40 Meet Elliot. He's a Canadian Forrest Pup. Unus...  
 42 This is Bella. She had her first beach experie...  
 43 Meet Jesse. He's a Fetty Woof. His tongue ejec...  
 45 This is Bailey. He thinks you should measure e...  
 47 This is Jack AKA Stephen Furry. You're not sco...  
 54 This is Penny. She's both pupset and fired pup...  
 56 This is Nelly. He graduated with his dogtorate...  
 57 This is Ginger. She's having a ruff Monday. To...  
 62 Meet Cash. He hath acquired a stick. A very go...  
 65 This is Sierra. She's one precious pupper. Abs...  
 68 This is Kody. He's a baller. Wishes he was a l...  
 70 This is Rover. As part of pupper protocol he h...  
 71 This is Napoleon. He's a Raggedy East Nicaragu...

79 This is Scout. He just graduated. Officially a...  
 82 Say hello to Cooper. His expression is the sam...  
 84 This is Shikha. She just watched you drop a sk...  
 85 This is Lili. She can't believe you betrayed h...  
 87 This is Coco. At first I thought she was a clo...  
 ...  
 972 This is Sweets the English Bulldog. Waves back...  
 995 This is Jimothy. He's a Trinidad Poliwirl. Fa...  
 998 This is Toffee. He's a happy pupper. Appears d...  
 1000 This is Carly. She's actually 2 dogs fused tog...  
 1013 This is Ellie. She's secretly ferocious. 12/10...  
 1017 Meet Percy. He's a Latvian Yuletide Heineken. ...  
 1019 I hope everyone enjoys this picture as much as...  
 1020 Meet Goliath. He's an example of irony. Head i...  
 1055 This is Reese. He likes holding hands. 12/10 h...  
 1082 This is Lenny. He was just told that he couldn...  
 1086 This is Daisy. She loves that shoe. Still no s...  
 1094 Meet Laela. She's adorable. Magnificent eyes. ...  
 1103 This is Ava. She doesn't understand flowers. 1...  
 1117 This is Kaiya. She's an aspiring shoe model. 1...  
 1121 This is Riley. She's just an adorable football...  
 1128 This is Kendall. 12/10 would cuddle the hell o...  
 1142 Say hello to Kenzie. She is a fluff ball. 12/1...  
 1161 Say hello to Jazz. She should be on the cover ...  
 1163 This is Franq and Pablo. They're working hard ...  
 1167 Meet Snickers. He's adorable. Also comes in t-...  
 1208 This is Remington. He's a man dime. 12/10 http...  
 1230 Meet Herb. 12/10 <https://t.co/tLRyYvCci3>  
 1232 Meet Scooter. He's ready for his first day of ...  
 1236 Meet Daisy. She's rebellious. Full of teen ang...  
 1254 This is Penny. She's having fun AND being safe...  
 1267 This is Kirk. He just saw a bacon wrapped tenn...  
 1272 This is Tucker. He is 100% ready for the sport...  
 1277 This is Chaz. He's an X Games half pipe supers...  
 1292 This is Gôrdôn. He enjoys his razberrita by po...  
 1300 This is Skittles. I would kidnap Skittles. Pin...

	expanded_urls	rating_numerator \
2	<a href="https://twitter.com/dog_rates/status/891815181...">https://twitter.com/dog_rates/status/891815181...</a>	12
4	<a href="https://twitter.com/dog_rates/status/891327558...">https://twitter.com/dog_rates/status/891327558...</a>	12
10	<a href="https://twitter.com/dog_rates/status/889638837...">https://twitter.com/dog_rates/status/889638837...</a>	12
13	<a href="https://twitter.com/dog_rates/status/888917238...">https://twitter.com/dog_rates/status/888917238...</a>	12
16	<a href="https://twitter.com/dog_rates/status/888078434...">https://twitter.com/dog_rates/status/888078434...</a>	12
22	<a href="https://twitter.com/dog_rates/status/886366144...">https://twitter.com/dog_rates/status/886366144...</a>	12
24	<a href="https://twitter.com/dog_rates/status/885984800...">https://twitter.com/dog_rates/status/885984800...</a>	12
26	<a href="https://twitter.com/dog_rates/status/884925521...">https://twitter.com/dog_rates/status/884925521...</a>	12
29	<a href="https://twitter.com/dog_rates/status/884162670...">https://twitter.com/dog_rates/status/884162670...</a>	12
30	<a href="https://twitter.com/dog_rates/status/883838122...">https://twitter.com/dog_rates/status/883838122...</a>	12

34	<a href="https://twitter.com/dog_rates/status/882762694...">https://twitter.com/dog_rates/status/882762694...</a>	12
38	<a href="https://twitter.com/dog_rates/status/881906580...">https://twitter.com/dog_rates/status/881906580...</a>	12
40	<a href="https://twitter.com/dog_rates/status/881268444...">https://twitter.com/dog_rates/status/881268444...</a>	12
42	<a href="https://twitter.com/dog_rates/status/880465832...">https://twitter.com/dog_rates/status/880465832...</a>	12
43	<a href="https://twitter.com/dog_rates/status/880221127...">https://twitter.com/dog_rates/status/880221127...</a>	12
45	<a href="https://twitter.com/dog_rates/status/879492040...">https://twitter.com/dog_rates/status/879492040...</a>	12
47	<a href="https://twitter.com/dog_rates/status/879376492...">https://twitter.com/dog_rates/status/879376492...</a>	12
54	<a href="https://twitter.com/dog_rates/status/877556246...">https://twitter.com/dog_rates/status/877556246...</a>	12
56	<a href="https://twitter.com/dog_rates/status/877201837...">https://twitter.com/dog_rates/status/877201837...</a>	12
57	<a href="https://twitter.com/dog_rates/status/876838120...">https://twitter.com/dog_rates/status/876838120...</a>	12
62	<a href="https://twitter.com/dog_rates/status/874680097...">https://twitter.com/dog_rates/status/874680097...</a>	12
65	<a href="https://www.gofundme.com/help-my-baby-sierra-g...">https://www.gofundme.com/help-my-baby-sierra-g...</a>	12
68	<a href="https://twitter.com/dog_rates/status/872122724...">https://twitter.com/dog_rates/status/872122724...</a>	12
70	<a href="https://twitter.com/dog_rates/status/871762521...">https://twitter.com/dog_rates/status/871762521...</a>	12
71	<a href="https://twitter.com/dog_rates/status/871515927...">https://twitter.com/dog_rates/status/871515927...</a>	12
79	<a href="https://twitter.com/dog_rates/status/869596645...">https://twitter.com/dog_rates/status/869596645...</a>	12
82	<a href="https://www.gofundme.com/3ti3nps">https://www.gofundme.com/3ti3nps</a> , <a href="https://twitt...">https://twitt...</a>	12
84	<a href="https://twitter.com/dog_rates/status/867421006...">https://twitter.com/dog_rates/status/867421006...</a>	12
85	<a href="https://twitter.com/dog_rates/status/866686824...">https://twitter.com/dog_rates/status/866686824...</a>	12
87	<a href="https://twitter.com/dog_rates/status/866334964...">https://twitter.com/dog_rates/status/866334964...</a>	12
...	...	...
972	<a href="https://vine.co/v/ib2nT0EuuOI">https://vine.co/v/ib2nT0EuuOI</a>	12
995	<a href="https://twitter.com/dog_rates/status/681981167...">https://twitter.com/dog_rates/status/681981167...</a>	12
998	<a href="https://twitter.com/dog_rates/status/681654059...">https://twitter.com/dog_rates/status/681654059...</a>	12
1000	<a href="https://twitter.com/dog_rates/status/681523177...">https://twitter.com/dog_rates/status/681523177...</a>	12
1013	<a href="https://twitter.com/dog_rates/status/680836378...">https://twitter.com/dog_rates/status/680836378...</a>	12
1017	<a href="https://twitter.com/dog_rates/status/680497766...">https://twitter.com/dog_rates/status/680497766...</a>	12
1019	<a href="https://twitter.com/dog_rates/status/680206703...">https://twitter.com/dog_rates/status/680206703...</a>	12
1020	<a href="https://twitter.com/dog_rates/status/680161097...">https://twitter.com/dog_rates/status/680161097...</a>	12
1055	<a href="https://twitter.com/dog_rates/status/678021115...">https://twitter.com/dog_rates/status/678021115...</a>	12
1082	<a href="https://twitter.com/dog_rates/status/675710890...">https://twitter.com/dog_rates/status/675710890...</a>	12
1086	<a href="https://twitter.com/dog_rates/status/675362609...">https://twitter.com/dog_rates/status/675362609...</a>	12
1094	<a href="https://twitter.com/dog_rates/status/675003128...">https://twitter.com/dog_rates/status/675003128...</a>	12
1103	<a href="https://twitter.com/dog_rates/status/674422304...">https://twitter.com/dog_rates/status/674422304...</a>	12
1117	<a href="https://twitter.com/dog_rates/status/674038233...">https://twitter.com/dog_rates/status/674038233...</a>	12
1121	<a href="https://twitter.com/dog_rates/status/673708611...">https://twitter.com/dog_rates/status/673708611...</a>	12
1128	<a href="https://twitter.com/dog_rates/status/673662677...">https://twitter.com/dog_rates/status/673662677...</a>	12
1142	<a href="https://twitter.com/dog_rates/status/672997845...">https://twitter.com/dog_rates/status/672997845...</a>	12
1161	<a href="https://twitter.com/dog_rates/status/672481316...">https://twitter.com/dog_rates/status/672481316...</a>	12
1163	<a href="https://twitter.com/dog_rates/status/672466075...">https://twitter.com/dog_rates/status/672466075...</a>	12
1167	<a href="https://twitter.com/dog_rates/status/672245253...">https://twitter.com/dog_rates/status/672245253...</a>	12
1208	<a href="https://twitter.com/dog_rates/status/670823764...">https://twitter.com/dog_rates/status/670823764...</a>	12
1230	<a href="https://twitter.com/dog_rates/status/670421925...">https://twitter.com/dog_rates/status/670421925...</a>	12
1232	<a href="https://twitter.com/dog_rates/status/670411370...">https://twitter.com/dog_rates/status/670411370...</a>	12
1236	<a href="https://twitter.com/dog_rates/status/670374371...">https://twitter.com/dog_rates/status/670374371...</a>	12
1254	<a href="https://twitter.com/dog_rates/status/669625907...">https://twitter.com/dog_rates/status/669625907...</a>	12
1267	<a href="https://twitter.com/dog_rates/status/669328503...">https://twitter.com/dog_rates/status/669328503...</a>	12
1272	<a href="https://twitter.com/dog_rates/status/669006782...">https://twitter.com/dog_rates/status/669006782...</a>	12

1277	<a href="https://twitter.com/dog_rates/status/668979806...">https://twitter.com/dog_rates/status/668979806...</a>	12
1292	<a href="https://twitter.com/dog_rates/status/668528771...">https://twitter.com/dog_rates/status/668528771...</a>	12
1300	<a href="https://twitter.com/dog_rates/status/668190681...">https://twitter.com/dog_rates/status/668190681...</a>	12

	rating_denominator	name	doggo	floofer	...	p1_conf \
2	10	Archie	None	None	...	0.716012
4	10	Franklin	None	None	...	0.555712
10	10	Ted	None	None	...	0.991650
13	10	Jim	None	None	...	0.714719
16	10	Gerald	None	None	...	0.995026
22	10	Roscoe	None	None	...	0.999201
24	10	Jimbo	None	None	...	0.972494
26	10	Earl	None	None	...	0.259916
29	10	Yogi	doggo	None	...	0.707046
30	10	Noah	None	None	...	0.610946
34	10	Gus	None	None	...	0.850050
38	10	Rey	None	None	...	0.291539
40	10	Elliot	None	None	...	0.473303
42	10	Bella	None	None	...	0.913255
43	10	Jesse	None	None	...	0.238525
45	10	Bailey	None	None	...	0.479896
47	10	Jack	None	None	...	0.663601
54	10	Penny	None	None	...	0.995368
56	10	Nelly	None	None	...	0.931120
57	10	Ginger	None	None	...	0.575751
62	10	Cash	None	None	...	0.836052
65	10	Sierra	None	None	...	0.619782
68	10	Kody	None	None	...	0.808396
70	10	Rover	None	None	...	0.921393
71	10	Napolean	doggo	None	...	0.974781
79	10	Scout	doggo	None	...	0.955156
82	10	Cooper	None	None	...	0.378151
84	10	Shikha	None	None	...	0.616457
85	10	Lili	None	None	...	0.514730
87	10	Coco	None	None	...	0.984086
...	...	...	...	...	...	...
972	10	Sweets	None	None	...	NaN
995	10	Jimothy	None	None	...	0.452577
998	10	Toffee	None	None	...	0.800538
1000	10	Carly	None	None	...	0.205067
1013	10	Ellie	None	None	...	0.427781
1017	10	Percy	None	None	...	0.538354
1019	10	Toby	None	None	...	0.149758
1020	10	Goliath	None	None	...	0.268681
1055	10	Reese	None	None	...	0.822048
1082	10	Lenny	None	None	...	0.441427
1086	10	Daisy	None	None	...	0.479008
1094	10	Laela	None	None	...	0.655279

1103	10	Ava	None	None	...	0.964497
1117	10	Kaiya	None	None	...	0.358459
1121	10	Riley	None	None	...	0.936333
1128	10	Kendall	None	None	...	0.957670
1142	10	Kenzie	None	None	...	0.517255
1161	10	Jazz	None	None	...	0.599454
1163	10	Franq	None	None	...	0.150424
1167	10	Snickers	None	None	...	0.718944
1208	10	Remington	None	None	...	0.947453
1230	10	Herb	None	None	...	0.275793
1232	10	Scooter	None	None	...	0.584397
1236	10	Daisy	None	None	...	0.974936
1254	10	Penny	None	None	...	0.874502
1267	10	Kirk	None	None	...	0.424202
1272	10	Tucker	None	None	...	0.127178
1277	10	Chaz	None	None	...	0.608537
1292	10	Gòrdón	None	None	...	0.195835
1300	10	Skittles	None	None	...	0.958402

	p1_dog	p2	p2_conf	p2_dog	\
2	True	Malamute	0.078253	True	
4	True	English_Springer	0.225770	True	
10	True	Boxer	0.002129	True	
13	True	Tibetan_Mastiff	0.120184	True	
16	True	Pug	0.000932	True	
22	True	Chihuahua	0.000361	True	
24	True	Shih-Tzu	0.006630	True	
26	True	American_Staffordshire_Terrier	0.198451	True	
29	True	Malinois	0.199396	True	
30	True	Miniature_Pinscher	0.299603	True	
34	True	Chesapeake_Bay_Retriever	0.074257	True	
38	True	Chesapeake_Bay_Retriever	0.278966	True	
40	False	Indian_Elephant	0.245646	False	
42	True	Labrador_Retriever	0.026329	True	
43	True	Meerkat	0.104256	False	
45	True	Vizsla	0.124353	True	
47	False	Labrador_Retriever	0.033496	True	
54	True	Welsh_Springer_Spaniel	0.001936	True	
56	True	Cardigan	0.068698	True	
57	True	Redbone	0.240970	True	
62	True	Staffordshire_Bullterrier	0.047069	True	
65	True	Bloodhound	0.338069	True	
68	False	Pug	0.066736	True	
70	True	Golden_Retriever	0.064608	True	
71	True	Briard	0.020041	True	
79	True	Toy_Terrier	0.008054	True	
82	True	Italian_Greyhound	0.275935	True	
84	True	Siberian_Husky	0.381330	True	

85	True	Groenendael	0.306407	True
87	True	Pomeranian	0.007919	True
...	...	...	...	...
972	NaN	NaN	NaN	NaN
995	True	Golden_Retriever	0.403420	True
998	True	Chow	0.146892	True
1000	True	German_Shepherd	0.160439	True
1013	True	Shetland_Sheepdog	0.160669	True
1017	True	Muzzle	0.084289	False
1019	False	Cloak	0.128830	False
1020	True	Miniature_Pinscher	0.125652	True
1055	True	Doberman	0.096085	True
1082	True	Miniature_Schnauzer	0.248885	True
1086	True	Ice_Bear	0.218289	False
1094	True	Pomeranian	0.104164	True
1103	True	Labrador_Retriever	0.009006	True
1117	True	Norwegian_Elkhound	0.206963	True
1121	True	Cocker_Spaniel	0.024211	True
1128	True	Beagle	0.012413	True
1142	True	Pomeranian	0.206053	True
1161	True	Collie	0.106227	True
1163	True	Toy_Poodle	0.088605	True
1167	True	Badger	0.178546	False
1208	True	German_Short-Haired_Pointer	0.017001	True
1230	True	Corn	0.073596	False
1232	True	Miniature_Schnauzer	0.064201	True
1236	True	English_Setter	0.011661	True
1254	False	Golden_Retriever	0.055408	True
1267	True	Eskimo_Dog	0.237660	True
1272	True	Italian_Greyhound	0.054215	True
1277	True	Irish_Setter	0.097078	True
1292	True	Kuvasz	0.121607	True
1300	True	Cocker_Spaniel	0.026764	True

	p3	p3_conf	p3_dog	retweet_count \
2	Kelpie	0.031379	True	4156.0
4	German_Short-Haired_Pointer	0.175219	True	9411.0
10	Staffordshire_Bullterrier	0.001498	True	4548.0
13	Labrador_Retriever	0.105506	True	4501.0
16	Bull_Mastiff	0.000903	True	3485.0
22	Boston_Bull	0.000076	True	3196.0
24	Bernese_Mountain_Dog	0.006239	True	6810.0
26	Staffordshire_Bullterrier	0.127725	True	18426.0
29	Norwegian_Elkhound	0.049148	True	2982.0
30	Kelpie	0.063020	True	3439.0
34	Flat-Coated_Retriever	0.015579	True	4918.0
38	Koala	0.127017	False	3398.0
40	Ibex	0.055661	False	5164.0



42	Cocker_Spaniel	0.009371	True	6283.0
43	Clumber	0.052580	True	4236.0
45	Bath_Towel	0.073320	False	3189.0
47	Pembroke	0.018827	True	3123.0
54	Bathtub	0.000468	False	3831.0
56	Basenji	0.000082	True	5616.0
57	Tibetan_Mastiff	0.088935	True	3366.0
62	Beagle	0.036007	True	4673.0
65	Chesapeake_Bay_Retriever	0.012676	True	1598.0
68	Dalmatian	0.054570	True	8203.0
70	Bloodhound	0.003383	True	3514.0
71	Swab	0.003228	False	3494.0
79	Muzzle	0.006296	False	3189.0
82	American_Staffordshire_Terrier	0.094991	True	2142.0
84	Malamute	0.001670	True	2576.0
85	Curly-Coated_Retriever	0.061314	True	3576.0
87	Keeshond	0.003328	True	14848.0
...	...	...	...	...
972	NaN	NaN	NaN	1631.0
995	Beagle	0.069486	True	1127.0
998	Pekinese	0.037613	True	1013.0
1000	Chow	0.156234	True	6359.0
1013	Pomeranian	0.111250	True	1453.0
1017	Ski_Mask	0.076690	False	2081.0
1019	Teddy	0.109129	False	1279.0
1020	English_Setter	0.089373	True	798.0
1055	Rottweiler	0.032709	True	6900.0
1082	Sealyham_Terrier	0.164967	True	893.0
1086	Kuvasz	0.139911	True	244.0
1094	Cardigan	0.052818	True	502.0
1103	Tennis_Ball	0.007139	False	563.0
1117	Malamute	0.148236	True	434.0
1121	Labrador_Retriever	0.009435	True	307.0
1128	Golden_Retriever	0.005689	True	382.0
1142	Koala	0.127037	False	759.0
1161	Shetland_Sheepdog	0.094465	True	128.0
1163	Welsh_Springer_Spaniel	0.072014	True	568.0
1167	Toy_Terrier	0.037103	True	163.0
1208	Weimaraner	0.015432	True	199.0
1230	Bolete	0.054905	False	672.0
1232	Toy_Poodle	0.060868	True	949.0
1236	Cocker_Spaniel	0.002689	True	280.0
1254	Labrador_Retriever	0.026854	True	1899.0
1267	Malamute	0.052572	True	436.0
1272	Pillow	0.048592	False	261.0
1277	Redbone	0.076022	True	371.0
1292	English_Setter	0.081464	True	235.0
1300	Welsh_Springer_Spaniel	0.007790	True	203.0

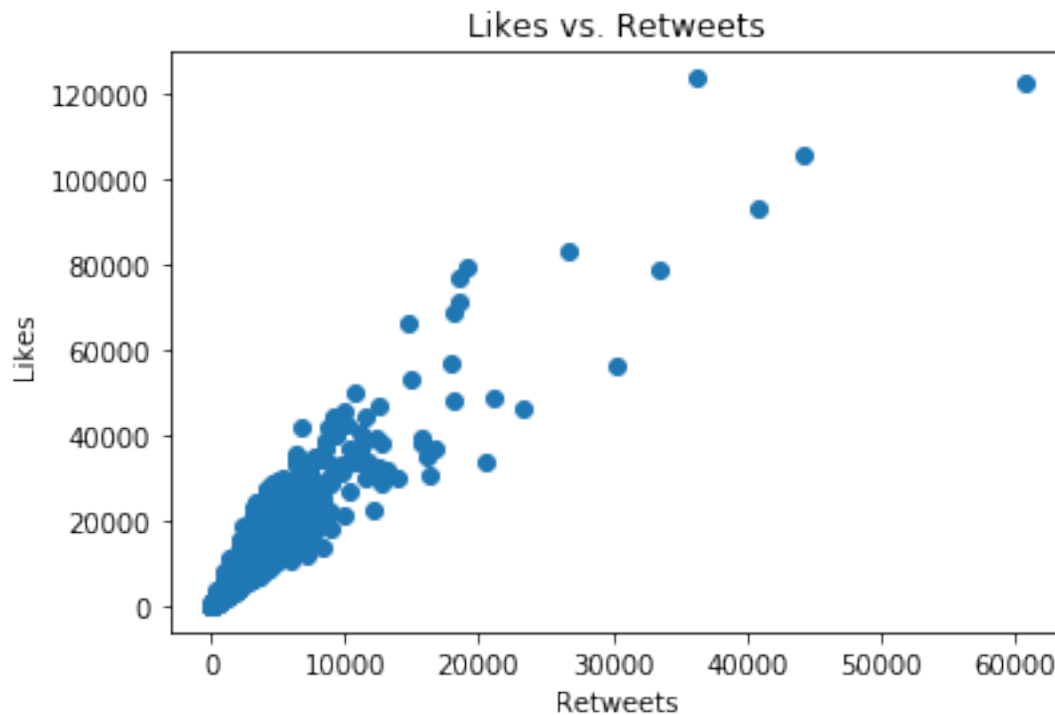
	favorite_count
2	24912.0
4	40140.0
10	27046.0
13	28948.0
16	21665.0
22	21096.0
24	32550.0
26	77188.0
29	20291.0
30	21807.0
34	28197.0
38	24204.0
40	22979.0
42	28397.0
43	26971.0
45	23314.0
47	16698.0
54	22674.0
56	27028.0
57	20602.0
62	27770.0
65	7238.0
68	34470.0
70	20287.0
71	20242.0
79	16070.0
82	10240.0
84	16303.0
85	19552.0
87	53257.0
...	...
972	3170.0
995	2955.0
998	2813.0
1000	15265.0
1013	3671.0
1017	4506.0
1019	3011.0
1020	2462.0
1055	14550.0
1082	1983.0
1086	1095.0
1094	1615.0
1103	1486.0
1117	1138.0
1121	1111.0

1128	1297.0
1142	2027.0
1161	736.0
1163	1401.0
1167	707.0
1208	799.0
1230	1355.0
1232	2105.0
1236	765.0
1254	3655.0
1267	1041.0
1272	595.0
1277	820.0
1292	479.0
1300	666.0

[333 rows x 25 columns]

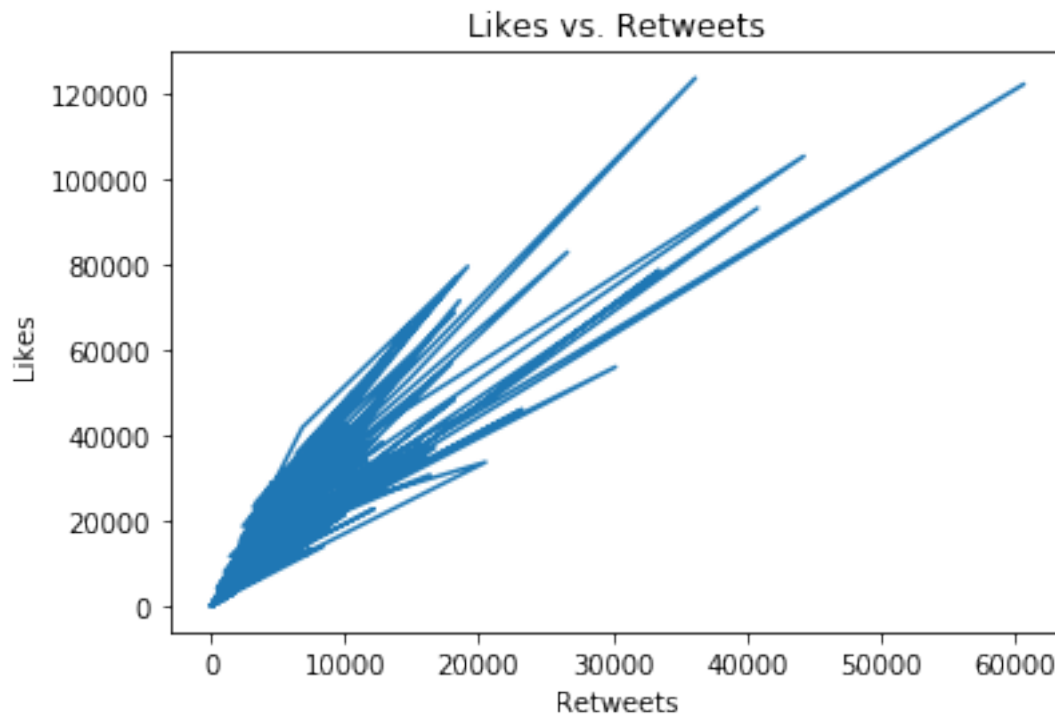
```
In [67]: #create scatter plot between retweets and favorites
plt.scatter(df_master['retweet_count'], df_master['favorite_count']);
plt.title('Likes vs. Retweets')
plt.xlabel('Retweets')
plt.ylabel('Likes')
```

```
Out[67]: Text(0,0.5,'Likes')
```



```
In [68]: #create scatter plot between retweets and favorites
plt.plot(df_master['retweet_count'], df_master['favorite_count']);
plt.title('Likes vs. Retweets')
plt.xlabel('Retweets')
plt.ylabel('Likes')
```

```
Out[68]: Text(0,0.5,'Likes')
```



**5.1.3 There is a psitive correlation between likes and retweets**