

# wrangle\_act

February 13, 2021

## 1 1 Twitter WeRateDogs Project

### 1.1 1.1 Gathering

#### 1.1.1 1.1.1 Gathering - twitter-archive-enhanced

```
In [1]: import pandas as pd
import numpy as np
import json
import re
```

```
df_1 = pd.read_csv('twitter-archive-enhanced.csv')
```

```
df_1
```

```
Out[1]:
```

	tweet_id	in_reply_to_status_id	in_reply_to_user_id	\
0	892420643555336193	NaN	NaN	
1	892177421306343426	NaN	NaN	
2	891815181378084864	NaN	NaN	
3	891689557279858688	NaN	NaN	
4	891327558926688256	NaN	NaN	
5	891087950875897856	NaN	NaN	
6	890971913173991426	NaN	NaN	
7	890729181411237888	NaN	NaN	
8	890609185150312448	NaN	NaN	
9	890240255349198849	NaN	NaN	
10	890006608113172480	NaN	NaN	
11	889880896479866881	NaN	NaN	
12	889665388333682689	NaN	NaN	
13	889638837579907072	NaN	NaN	
14	889531135344209921	NaN	NaN	
15	889278841981685760	NaN	NaN	
16	888917238123831296	NaN	NaN	
17	888804989199671297	NaN	NaN	
18	888554962724278272	NaN	NaN	
19	888202515573088257	NaN	NaN	
20	888078434458587136	NaN	NaN	
21	887705289381826560	NaN	NaN	

22	887517139158093824	NaN	NaN
23	887473957103951883	NaN	NaN
24	887343217045368832	NaN	NaN
25	887101392804085760	NaN	NaN
26	886983233522544640	NaN	NaN
27	886736880519319552	NaN	NaN
28	886680336477933568	NaN	NaN
29	886366144734445568	NaN	NaN
...	...	...	...
2326	666411507551481857	NaN	NaN
2327	666407126856765440	NaN	NaN
2328	666396247373291520	NaN	NaN
2329	666373753744588802	NaN	NaN
2330	666362758909284353	NaN	NaN
2331	666353288456101888	NaN	NaN
2332	666345417576210432	NaN	NaN
2333	666337882303524864	NaN	NaN
2334	666293911632134144	NaN	NaN
2335	666287406224695296	NaN	NaN
2336	666273097616637952	NaN	NaN
2337	666268910803644416	NaN	NaN
2338	666104133288665088	NaN	NaN
2339	666102155909144576	NaN	NaN
2340	666099513787052032	NaN	NaN
2341	666094000022159362	NaN	NaN
2342	666082916733198337	NaN	NaN
2343	666073100786774016	NaN	NaN
2344	666071193221509120	NaN	NaN
2345	666063827256086533	NaN	NaN
2346	666058600524156928	NaN	NaN
2347	666057090499244032	NaN	NaN
2348	666055525042405380	NaN	NaN
2349	666051853826850816	NaN	NaN
2350	666050758794694657	NaN	NaN
2351	666049248165822465	NaN	NaN
2352	666044226329800704	NaN	NaN
2353	666033412701032449	NaN	NaN
2354	666029285002620928	NaN	NaN
2355	666020888022790149	NaN	NaN

	timestamp \
0	2017-08-01 16:23:56 +0000
1	2017-08-01 00:17:27 +0000
2	2017-07-31 00:18:03 +0000
3	2017-07-30 15:58:51 +0000
4	2017-07-29 16:00:24 +0000
5	2017-07-29 00:08:17 +0000
6	2017-07-28 16:27:12 +0000

7	2017-07-28	00:22:40	+0000
8	2017-07-27	16:25:51	+0000
9	2017-07-26	15:59:51	+0000
10	2017-07-26	00:31:25	+0000
11	2017-07-25	16:11:53	+0000
12	2017-07-25	01:55:32	+0000
13	2017-07-25	00:10:02	+0000
14	2017-07-24	17:02:04	+0000
15	2017-07-24	00:19:32	+0000
16	2017-07-23	00:22:39	+0000
17	2017-07-22	16:56:37	+0000
18	2017-07-22	00:23:06	+0000
19	2017-07-21	01:02:36	+0000
20	2017-07-20	16:49:33	+0000
21	2017-07-19	16:06:48	+0000
22	2017-07-19	03:39:09	+0000
23	2017-07-19	00:47:34	+0000
24	2017-07-18	16:08:03	+0000
25	2017-07-18	00:07:08	+0000
26	2017-07-17	16:17:36	+0000
27	2017-07-16	23:58:41	+0000
28	2017-07-16	20:14:00	+0000
29	2017-07-15	23:25:31	+0000
...			...
2326	2015-11-17	00:24:19	+0000
2327	2015-11-17	00:06:54	+0000
2328	2015-11-16	23:23:41	+0000
2329	2015-11-16	21:54:18	+0000
2330	2015-11-16	21:10:36	+0000
2331	2015-11-16	20:32:58	+0000
2332	2015-11-16	20:01:42	+0000
2333	2015-11-16	19:31:45	+0000
2334	2015-11-16	16:37:02	+0000
2335	2015-11-16	16:11:11	+0000
2336	2015-11-16	15:14:19	+0000
2337	2015-11-16	14:57:41	+0000
2338	2015-11-16	04:02:55	+0000
2339	2015-11-16	03:55:04	+0000
2340	2015-11-16	03:44:34	+0000
2341	2015-11-16	03:22:39	+0000
2342	2015-11-16	02:38:37	+0000
2343	2015-11-16	01:59:36	+0000
2344	2015-11-16	01:52:02	+0000
2345	2015-11-16	01:22:45	+0000
2346	2015-11-16	01:01:59	+0000
2347	2015-11-16	00:55:59	+0000
2348	2015-11-16	00:49:46	+0000
2349	2015-11-16	00:35:11	+0000

2350 2015-11-16 00:30:50 +0000  
 2351 2015-11-16 00:24:50 +0000  
 2352 2015-11-16 00:04:52 +0000  
 2353 2015-11-15 23:21:54 +0000  
 2354 2015-11-15 23:05:30 +0000  
 2355 2015-11-15 22:32:08 +0000

```

                                source \
0      <a href="http://twitter.com/download/iphone" r...
1      <a href="http://twitter.com/download/iphone" r...
2      <a href="http://twitter.com/download/iphone" r...
3      <a href="http://twitter.com/download/iphone" r...
4      <a href="http://twitter.com/download/iphone" r...
5      <a href="http://twitter.com/download/iphone" r...
6      <a href="http://twitter.com/download/iphone" r...
7      <a href="http://twitter.com/download/iphone" r...
8      <a href="http://twitter.com/download/iphone" r...
9      <a href="http://twitter.com/download/iphone" r...
10     <a href="http://twitter.com/download/iphone" r...
11     <a href="http://twitter.com/download/iphone" r...
12     <a href="http://twitter.com/download/iphone" r...
13     <a href="http://twitter.com/download/iphone" r...
14     <a href="http://twitter.com/download/iphone" r...
15     <a href="http://twitter.com/download/iphone" r...
16     <a href="http://twitter.com/download/iphone" r...
17     <a href="http://twitter.com/download/iphone" r...
18     <a href="http://twitter.com/download/iphone" r...
19     <a href="http://twitter.com/download/iphone" r...
20     <a href="http://twitter.com/download/iphone" r...
21     <a href="http://twitter.com/download/iphone" r...
22     <a href="http://twitter.com/download/iphone" r...
23     <a href="http://twitter.com/download/iphone" r...
24     <a href="http://twitter.com/download/iphone" r...
25     <a href="http://twitter.com/download/iphone" r...
26     <a href="http://twitter.com/download/iphone" r...
27     <a href="http://twitter.com/download/iphone" r...
28     <a href="http://twitter.com/download/iphone" r...
29     <a href="http://twitter.com/download/iphone" r...
...
2326  <a href="http://twitter.com/download/iphone" r...
2327  <a href="http://twitter.com/download/iphone" r...
2328  <a href="http://twitter.com/download/iphone" r...
2329  <a href="http://twitter.com/download/iphone" r...
2330  <a href="http://twitter.com/download/iphone" r...
2331  <a href="http://twitter.com/download/iphone" r...
2332  <a href="http://twitter.com/download/iphone" r...
2333  <a href="http://twitter.com/download/iphone" r...
2334  <a href="http://twitter.com/download/iphone" r...

```

2335 <a href="http://twitter.com/download/iphone" r...  
 2336 <a href="http://twitter.com/download/iphone" r...  
 2337 <a href="http://twitter.com/download/iphone" r...  
 2338 <a href="http://twitter.com/download/iphone" r...  
 2339 <a href="http://twitter.com/download/iphone" r...  
 2340 <a href="http://twitter.com/download/iphone" r...  
 2341 <a href="http://twitter.com/download/iphone" r...  
 2342 <a href="http://twitter.com/download/iphone" r...  
 2343 <a href="http://twitter.com/download/iphone" r...  
 2344 <a href="http://twitter.com/download/iphone" r...  
 2345 <a href="http://twitter.com/download/iphone" r...  
 2346 <a href="http://twitter.com/download/iphone" r...  
 2347 <a href="http://twitter.com/download/iphone" r...  
 2348 <a href="http://twitter.com/download/iphone" r...  
 2349 <a href="http://twitter.com/download/iphone" r...  
 2350 <a href="http://twitter.com/download/iphone" r...  
 2351 <a href="http://twitter.com/download/iphone" r...  
 2352 <a href="http://twitter.com/download/iphone" r...  
 2353 <a href="http://twitter.com/download/iphone" r...  
 2354 <a href="http://twitter.com/download/iphone" r...  
 2355 <a href="http://twitter.com/download/iphone" r...

	text	retweeted_status_id \
0	This is Phineas. He's a mystical boy. Only eve...	NaN
1	This is Tilly. She's just checking pup on you...	NaN
2	This is Archie. He is a rare Norwegian Pouncin...	NaN
3	This is Darla. She commenced a snooze mid meal...	NaN
4	This is Franklin. He would like you to stop ca...	NaN
5	Here we have a majestic great white breaching ...	NaN
6	Meet Jax. He enjoys ice cream so much he gets ...	NaN
7	When you watch your owner call another dog a g...	NaN
8	This is Zoey. She doesn't want to be one of th...	NaN
9	This is Cassie. She is a college pup. Studying...	NaN
10	This is Koda. He is a South Australian decksha...	NaN
11	This is Bruno. He is a service shark. Only get...	NaN
12	Here's a puppo that seems to be on the fence a...	NaN
13	This is Ted. He does his best. Sometimes that'...	NaN
14	This is Stuart. He's sporting his favorite fan...	NaN
15	This is Oliver. You're witnessing one of his m...	NaN
16	This is Jim. He found a fren. Taught him how t...	NaN
17	This is Zeke. He has a new stick. Very proud o...	NaN
18	This is Ralphus. He's powering up. Attempting ...	NaN
19	RT @dog_rates: This is Canela. She attempted s...	8.874740e+17
20	This is Gerald. He was just told he didn't get...	NaN
21	This is Jeffrey. He has a monopoly on the pool...	NaN
22	I've yet to rate a Venezuelan Hover Wiener. Th...	NaN
23	This is Canela. She attempted some fancy porch...	NaN
24	You may not have known you needed to see this ...	NaN

25	This... is a Jubilant Antarctic House Bear. We...	NaN
26	This is Maya. She's very shy. Rarely leaves he...	NaN
27	This is Mingus. He's a wonderful father to his...	NaN
28	This is Derek. He's late for a dog meeting. 13...	NaN
29	This is Roscoe. Another pupper fallen victim t...	NaN
...	...	...
2326	This is quite the dog. Gets really excited whe...	NaN
2327	This is a southern Vesuvius bumblegruff. Can d...	NaN
2328	Oh goodness. A super rare northeast Qdoba kang...	NaN
2329	Those are sunglasses and a jean jacket. 11/10 ...	NaN
2330	Unique dog here. Very small. Lives in containe...	NaN
2331	Here we have a mixed Asiago from the Galápagos...	NaN
2332	Look at this jokester thinking seat belt laws ...	NaN
2333	This is an extremely rare horned Parthenon. No...	NaN
2334	This is a funny dog. Weird toes. Won't come do...	NaN
2335	This is an Albanian 3 1/2 legged Episcopalian...	NaN
2336	Can take selfies 11/10 <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	https://twitter.com/dog_rates/status/892420643...	13
1	https://twitter.com/dog_rates/status/892177421...	13
2	https://twitter.com/dog_rates/status/891815181...	12
3	https://twitter.com/dog_rates/status/891689557...	13
4	https://twitter.com/dog_rates/status/891327558...	12
5	https://twitter.com/dog_rates/status/891087950...	13
6	https://gofundme.com/ydvmve-surgery-for-jax,ht...	13
7	https://twitter.com/dog_rates/status/890729181...	13
8	https://twitter.com/dog_rates/status/890609185...	13
9	https://twitter.com/dog_rates/status/890240255...	14
10	https://twitter.com/dog_rates/status/890006608...	13
11	https://twitter.com/dog_rates/status/889880896...	13
12	https://twitter.com/dog_rates/status/889665388...	13
13	https://twitter.com/dog_rates/status/889638837...	12
14	https://twitter.com/dog_rates/status/889531135...	13
15	https://twitter.com/dog_rates/status/889278841...	13
16	https://twitter.com/dog_rates/status/888917238...	12
17	https://twitter.com/dog_rates/status/888804989...	13
18	https://twitter.com/dog_rates/status/888554962...	13
19	https://twitter.com/dog_rates/status/887473957...	13
20	https://twitter.com/dog_rates/status/888078434...	12
21	https://twitter.com/dog_rates/status/887705289...	13
22	https://twitter.com/dog_rates/status/887517139...	14
23	https://twitter.com/dog_rates/status/887473957...	13
24	https://twitter.com/dog_rates/status/887343217...	13
25	https://twitter.com/dog_rates/status/887101392...	12
26	https://twitter.com/dog_rates/status/886983233...	13
27	https://www.gofundme.com/mingusneedsus,https:/...	13
28	https://twitter.com/dog_rates/status/886680336...	13
29	https://twitter.com/dog_rates/status/886366144...	12
...	...	...
2326	https://twitter.com/dog_rates/status/666411507...	2
2327	https://twitter.com/dog_rates/status/666407126...	7
2328	https://twitter.com/dog_rates/status/666396247...	9
2329	https://twitter.com/dog_rates/status/666373753...	11
2330	https://twitter.com/dog_rates/status/666362758...	6
2331	https://twitter.com/dog_rates/status/666353288...	8
2332	https://twitter.com/dog_rates/status/666345417...	10
2333	https://twitter.com/dog_rates/status/666337882...	9
2334	https://twitter.com/dog_rates/status/666293911...	3
2335	https://twitter.com/dog_rates/status/666287406...	1
2336	https://twitter.com/dog_rates/status/666273097...	11
2337	https://twitter.com/dog_rates/status/666268910...	10



2338	<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]

### 1.1.2 1.1.2 Gathering - image-predictions

```
In [2]: import pandas as pd
import requests
```

```
url = 'https://d17h27t6h515a5.cloudfront.net/topher/2017/August/599fd2ad_image-predictions'
r = requests.get(url, allow_redirects=True)
```

```
open('image-predictions.tsv', 'wb').write(r.content)
```

```
df_2 = pd.read_csv('image-predictions.tsv', sep='\t')
```

df\_2

```
Out[2]:
```

	tweet_id	jpg_url \
0	666020888022790149	<a href="https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg">https://pbs.twimg.com/media/CT4udn0WwAA0aMy.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/CT5w9gUW4AAAsBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAAsBNN.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/CT54YGiwUAEZnOk.jpg">https://pbs.twimg.com/media/CT54YGiwUAEZnOk.jpg</a>
17	666104133288665088	<a href="https://pbs.twimg.com/media/CT56LSZWwAA1Jj2.jpg">https://pbs.twimg.com/media/CT56LSZWwAA1Jj2.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/CT9Vn7PWwAA_ZCM.jpg">https://pbs.twimg.com/media/CT9Vn7PWwAA_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-NvwmW4AAAugGZ.jpg">https://pbs.twimg.com/media/CT-NvwmW4AAAugGZ.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/DE8yicJWOAAAyBJ.jpg">https://pbs.twimg.com/media/DE8yicJWOAAAyBJ.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-UQAACu20.jpg">https://pbs.twimg.com/media/DFTH_0-UQAACu20.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/DF1eOmZXUAAALUc.jpg">https://pbs.twimg.com/media/DF1eOmZXUAAALUc.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/DF6hr6BUMAAzZgT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.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	terrapien
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]

### 1.1.3 1.1.3 Gathering - tweepy.API

```
In [3]: df_list = []
```

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

df_3 = pd.DataFrame(df_list)
df_3.head()
```

```
Out[3]:
```

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

## 1.2 1.2 Assess

```
In [4]: df_1.sample(10)
```

```
Out[4]:
```

	tweet_id	in_reply_to_status_id	in_reply_to_user_id	\
398	825026590719483904	NaN	NaN	
2070	671122204919246848	NaN	NaN	
2082	670815497391357952	NaN	NaN	



755	778748913645780993	NaN	NaN
601	798665375516884993	NaN	NaN
610	797236660651966464	NaN	NaN
277	840370681858686976	NaN	NaN
912	757596066325864448	NaN	NaN
691	787322443945877504	NaN	NaN
1819	676590572941893632	6.765883e+17	4.196984e+09

	timestamp \
398	2017-01-27 17:04:02 +0000
2070	2015-11-30 00:22:57 +0000
2082	2015-11-29 04:04:12 +0000
755	2016-09-22 00:13:04 +0000
601	2016-11-15 23:13:58 +0000
610	2016-11-12 00:36:46 +0000
277	2017-03-11 01:15:58 +0000
912	2016-07-25 15:19:12 +0000
691	2016-10-15 16:01:13 +0000
1819	2015-12-15 02:32:17 +0000

	source \
398	<a href="http://twitter.com/download/iphone" r...
2070	<a href="http://twitter.com/download/iphone" r...
2082	<a href="http://twitter.com/download/iphone" r...
755	<a href="http://twitter.com/download/iphone" r...
601	<a href="http://twitter.com/download/iphone" r...
610	<a href="http://twitter.com/download/iphone" r...
277	<a href="http://twitter.com/download/iphone" r...
912	<a href="http://twitter.com/download/iphone" r...
691	<a href="http://twitter.com/download/iphone" r...
1819	<a href="http://twitter.com/download/iphone" r...

	text	retweeted_status_id \
398	Say hello to Pablo. He's one gorgeous puppo. A...	NaN
2070	Two miniature golden retrievers here. Webbed p...	NaN
2082	This is Sage. He likes to burn shit. 10/10 htt...	NaN
755	This is Mya (pronounced "mmmyah?"). Her head i...	NaN
601	RT @dog_rates: This is Lola. She fell asleep o...	6.718968e+17
610	This is Pancake. She loves Batman and winks li...	NaN
277	You have been visited by the magical sugar jar...	NaN
912	Here's another picture without a dog in it. Id...	NaN
691	This is Lincoln. He forgot to use his blinker ...	NaN
1819	After some outrage from the crowd. Bubbles is ...	NaN

	retweeted_status_user_id	retweeted_status_timestamp \
398	NaN	NaN
2070	NaN	NaN
2082	NaN	NaN

755	NaN	NaN
601	4.196984e+09	2015-12-02 03:40:57 +0000
610	NaN	NaN
277	NaN	NaN
912	NaN	NaN
691	NaN	NaN
1819	NaN	NaN

	expanded_urls	rating_numerator	\
398	https://www.gofundme.com/my-puppys-double-cata...	12	
2070	https://twitter.com/dog_rates/status/671122204...	4	
2082	https://twitter.com/dog_rates/status/670815497...	10	
755	https://twitter.com/dog_rates/status/778748913...	11	
601	https://twitter.com/dog_rates/status/671896809...	10	
610	https://twitter.com/dog_rates/status/797236660...	12	
277	https://twitter.com/dog_rates/status/840370681...	13	
912	https://twitter.com/dog_rates/status/757596066...	4	
691	https://twitter.com/dog_rates/status/787322443...	10	
1819	NaN	7	

	rating_denominator	name	doggo	floofer	pupper	puppo
398	10	Pablo	None	None	None	puppo
2070	10	None	None	None	None	None
2082	10	Sage	None	None	None	None
755	10	Mya	None	None	None	None
601	10	Lola	None	None	None	None
610	10	Pancake	None	None	None	None
277	10	None	None	None	None	None
912	10	None	None	None	None	None
691	10	Lincoln	None	None	None	None
1819	10	None	None	None	None	None

```
In [5]: df_1.shape
```

```
Out[5]: (2356, 17)
```

```
In [6]: df_1.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 [7]: df_1.describe()
```

```

Out[7]:
          tweet_id  in_reply_to_status_id  in_reply_to_user_id \
count    2.356000e+03          7.800000e+01          7.800000e+01
mean     7.427716e+17          7.455079e+17          2.014171e+16
std      6.856705e+16          7.582492e+16          1.252797e+17
min      6.660209e+17          6.658147e+17          1.185634e+07
25%      6.783989e+17          6.757419e+17          3.086374e+08
50%      7.196279e+17          7.038708e+17          4.196984e+09
75%      7.993373e+17          8.257804e+17          4.196984e+09
max      8.924206e+17          8.862664e+17          8.405479e+17

          retweeted_status_id  retweeted_status_user_id  rating_numerator \
count          1.810000e+02          1.810000e+02          2356.000000
mean          7.720400e+17          1.241698e+16          13.126486
std           6.236928e+16          9.599254e+16          45.876648
min           6.661041e+17          7.832140e+05           0.000000
25%           7.186315e+17          4.196984e+09          10.000000
50%           7.804657e+17          4.196984e+09          11.000000
75%           8.203146e+17          4.196984e+09          12.000000
max           8.874740e+17          7.874618e+17          1776.000000

          rating_denominator
count          2356.000000
mean           10.455433
std            6.745237
min            0.000000
25%           10.000000
50%           10.000000
75%           10.000000
max           170.000000

```

```
In [8]: df_1['tweet_id'].duplicated().sum()
```

```
Out[8]: 0
```

```
In [9]: df_1['source'].unique()
```

```
Out[9]: array([ '<a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone</a>',  
                '<a href="http://twitter.com" rel="nofollow">Twitter Web Client</a>',  
                '<a href="http://vine.co" rel="nofollow">Vine - Make a Scene</a>',  
                '<a href="https://about.twitter.com/products/tweetdeck" rel="nofollow">TweetDeck</a>'])
```

```
In [10]: with pd.option_context('display.max_colwidth', None):  
         display(df_1.text.head())
```

-----  
ValueError

Traceback (most recent call last)

```
<ipython-input-10-06136fa7cc2e> in <module>()
----> 1 with pd.option_context('display.max_colwidth', None):
      2     display(df_1.text.head())

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in __enter__(self)
404
405     for pat, val in self.ops:
--> 406         _set_option(pat, val, silent=True)
407
408     def __exit__(self, *args):

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in _set_option(*args, **kwargs)
121     o = _get_registered_option(key)
122     if o and o.validator:
--> 123         o.validator(v)
124
125     # walk the nested dict

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in inner(x)
762     if type(x) != _type:
763         msg = "Value must have type '{typ!s}'"
--> 764         raise ValueError(msg.format(typ=_type))
765
766     return inner
```

ValueError: Value must have type '<class 'int'>'

```
In [ ]: df_1['expanded_urls'].isnull().sum()
```

```

In [ ]: with pd.option_context('display.max_colwidth', None):
        display(df_1.expanded_urls.head(50))

In [ ]: df_1['rating_numerator'].value_counts()

In [ ]: x = df_1.query('rating_numerator <= 5')

        with pd.option_context('display.max_colwidth', None):
            display(x.sample(30))

In [ ]: y = df_1.query('rating_numerator >= 15')

        with pd.option_context('display.max_colwidth', None):
            display(y)

In [ ]: df_1['rating_denominator'].value_counts()

In [ ]: df_1['rating_denominator'][df_1['rating_denominator'] != 10].value_counts()

In [ ]: with pd.option_context('display.max_colwidth', None):
        display(df_1[df_1['rating_denominator'] != 10].text)

In [ ]: df_1['name'].value_counts()

In [11]: df_1['name'].unique()

Out[11]: array(['Phineas', 'Tilly', 'Archie', 'Darla', 'Franklin', 'None', 'Jax',
                'Zoey', 'Cassie', 'Koda', 'Bruno', 'Ted', 'Stuart', 'Oliver', 'Jim',
                'Zeke', 'Ralphus', 'Canela', 'Gerald', 'Jeffrey', 'such', 'Maya',
                'Mingus', 'Derek', 'Roscoe', 'Waffles', 'Jimbo', 'Maisey', 'Lilly',
                'Earl', 'Lola', 'Kevin', 'Yogi', 'Noah', 'Bella', 'Grizzwald',
                'Rusty', 'Gus', 'Stanley', 'Alfy', 'Koko', 'Rey', 'Gary', 'a',
                'Elliot', 'Louis', 'Jesse', 'Romeo', 'Bailey', 'Duddles', 'Jack',
                'Emmy', 'Steven', 'Beau', 'Snoopy', 'Shadow', 'Terrance', 'Aja',
                'Penny', 'Dante', 'Nelly', 'Ginger', 'Benedict', 'Venti', 'Goose',
                'Nugget', 'Cash', 'Coco', 'Jed', 'Sebastian', 'Walter', 'Sierra',
                'Monkey', 'Harry', 'Kody', 'Lassie', 'Rover', 'Napolean', 'Dawn',
                'Boomer', 'Cody', 'Rumble', 'Clifford', 'quite', 'Dewey', 'Scout',
                'Gizmo', 'Cooper', 'Harold', 'Shikha', 'Jamesy', 'Lili', 'Sammy',
                'Meatball', 'Paisley', 'Albus', 'Neptune', 'Quinn', 'Belle',
                'Zooey', 'Dave', 'Jersey', 'Hobbes', 'Burt', 'Lorenzo', 'Carl',
                'Jordy', 'Milky', 'Trooper', 'Winston', 'Sophie', 'Wyatt', 'Rosie',
                'Thor', 'Oscar', 'Luna', 'Callie', 'Cermet', 'George', 'Marlee',
                'Arya', 'Einstein', 'Alice', 'Rumpole', 'Benny', 'Aspen', 'Jarod',
                'Wiggles', 'General', 'Sailor', 'Astrid', 'Iggy', 'Snoop', 'Kyle',
                'Leo', 'Riley', 'Gidget', 'Noosh', 'Odin', 'Jerry', 'Charlie',
                'Georgie', 'Rontu', 'Cannon', 'Furzey', 'Daisy', 'Tuck', 'Barney',
                'Vixen', 'Jarvis', 'Mimosa', 'Pickles', 'Bungalo', 'Brady', 'Margo',
                'Sadie', 'Hank', 'Tycho', 'Stephan', 'Indie', 'Winnie', 'Bentley',

```

'Ken', 'Max', 'Maddie', 'Pipsy', 'Monty', 'Sojourner', 'Odie',  
 'Arlo', 'Sunny', 'Vincent', 'Lucy', 'Clark', 'Mookie', 'Meera',  
 'Buddy', 'Ava', 'Rory', 'Eli', 'Ash', 'Tucker', 'Tobi', 'Chester',  
 'Wilson', 'Sunshine', 'Lipton', 'Gabby', 'Bronte', 'Poppy', 'Rhino',  
 'Willow', 'not', 'Orion', 'Eevee', 'Smiley', 'Logan', 'Moreton',  
 'Klein', 'Miguel', 'Emanuel', 'Kuyu', 'Dutch', 'Pete', 'Scooter',  
 'Reggie', 'Kyro', 'Samson', 'Loki', 'Mia', 'Malcolm', 'Dexter',  
 'Alfie', 'Fiona', 'one', 'Mutt', 'Bear', 'Doobert', 'Beebop',  
 'Alexander', 'Sailer', 'Brutus', 'Kona', 'Boots', 'Ralphie', 'Phil',  
 'Cupid', 'Pawnd', 'Pilot', 'Ike', 'Mo', 'Toby', 'Sweet', 'Pablo',  
 'Nala', 'Balto', 'Crawford', 'Gabe', 'Mattie', 'Jimison',  
 'Hercules', 'Duchess', 'Harlso', 'Sampson', 'Sundance', 'Luca',  
 'Flash', 'Finn', 'Peaches', 'Howie', 'Jazzy', 'Anna', 'Bo',  
 'Seamus', 'Wafer', 'Chelsea', 'Tom', 'Moose', 'Florence', 'Autumn',  
 'Dido', 'Eugene', 'Herschel', 'Strudel', 'Tebow', 'Chloe', 'Betty',  
 'Timber', 'Binky', 'Dudley', 'Comet', 'Larry', 'Levi', 'Akumi',  
 'Titan', 'Olivia', 'Alf', 'Oshie', 'Bruce', 'Chubbs', 'Sky',  
 'Atlas', 'Eleanor', 'Layla', 'Rocky', 'Baron', 'Tyr', 'Bauer',  
 'Swagger', 'Brandi', 'Mary', 'Moe', 'Halo', 'Augie', 'Craig', 'Sam',  
 'Hunter', 'Pavlov', 'Maximus', 'Wallace', 'Ito', 'Milo', 'Ollie',  
 'Cali', 'Lennon', 'incredibly', 'Major', 'Duke', 'Reginald',  
 'Sansa', 'Shooter', 'Django', 'Diogi', 'Sonny', 'Philbert',  
 'Marley', 'Severus', 'Ronnie', 'Anakin', 'Bones', 'Mauve', 'Chef',  
 'Doc', 'Sobe', 'Longfellow', 'Mister', 'Iroh', 'Baloo', 'Stubert',  
 'Paull', 'Timison', 'Davey', 'Pancake', 'Tyrone', 'Snicku', 'Ruby',  
 'Brody', 'Rizzy', 'Mack', 'Butter', 'Nimbus', 'Laika', 'Dobby',  
 'Juno', 'Maude', 'Lily', 'Newt', 'Benji', 'Nida', 'Robin',  
 'Monster', 'BeBe', 'Remus', 'Mabel', 'Misty', 'Happy', 'Mosby',  
 'Maggie', 'Leela', 'Ralphie', 'Brownie', 'Meyer', 'Stella', 'mad',  
 'Frank', 'Tonks', 'Lincoln', 'Oakley', 'Dale', 'Rizzo', 'Arnie',  
 'Pinot', 'Dallas', 'Hero', 'Frankie', 'Stormy', 'Mairi', 'Loomis',  
 'Godi', 'Kenny', 'Deacon', 'Timmy', 'Harper', 'Chipson', 'Combo',  
 'Dash', 'Bell', 'Hurley', 'Jay', 'Mya', 'Strider', 'an', 'Wesley',  
 'Solomon', 'Huck', 'very', 'O', 'Blue', 'Finley', 'Sprinkles',  
 'Heinrich', 'Shakespeare', 'Fizz', 'Chip', 'Grey', 'Roosevelt',  
 'Gromit', 'Willem', 'Dakota', 'Dixie', 'Al', 'Jackson', 'just',  
 'Carbon', 'DonDon', 'Kirby', 'Lou', 'Nollie', 'Chevy', 'Tito',  
 'Louie', 'Rupert', 'Rufus', 'Brudge', 'Shadoe', 'Colby', 'Angel',  
 'Brat', 'Tove', 'my', 'Aubie', 'Kota', 'Eve', 'Glenn', 'Shelby',  
 'Sephie', 'Bonaparte', 'Albert', 'Wishes', 'Rose', 'Theo', 'Rocco',  
 'Fido', 'Emma', 'Spencer', 'Lilli', 'Boston', 'Brandonald', 'Corey',  
 'Leonard', 'Chompsky', 'Beckham', 'Devón', 'Gert', 'Watson',  
 'Rubio', 'Keith', 'Dex', 'Carly', 'Ace', 'Tayzie', 'Grizzie',  
 'Fred', 'Gilbert', 'Zoe', 'Stewie', 'Calvin', 'Lilah', 'Spanky',  
 'Jameson', 'Piper', 'Atticus', 'Blu', 'Dietrich', 'Divine', 'Tripp',  
 'his', 'Cora', 'Huxley', 'Keurig', 'Bookstore', 'Linus', 'Abby',  
 'Shaggy', 'Shiloh', 'Gustav', 'Arlen', 'Percy', 'Lenox', 'Sugar',  
 'Harvey', 'Blanket', 'actually', 'Geno', 'Stark', 'Beya', 'Kilo',

'Kayla', 'Maxaroni', 'Doug', 'Edmund', 'Aqua', 'Theodore', 'Chase',  
 'getting', 'Rorie', 'Simba', 'Charles', 'Bayley', 'Axel',  
 'Storkson', 'Remy', 'Chadrick', 'Kellogg', 'Buckley', 'Livvie',  
 'Terry', 'Hermione', 'Ralpher', 'Aldrick', 'this', 'unacceptable',  
 'Rooney', 'Crystal', 'Ziva', 'Stefan', 'Pupcasso', 'Puff',  
 'Flurpson', 'Coleman', 'Enchilada', 'Raymond', 'all', 'Rueben',  
 'Cilantro', 'Karll', 'Sprout', 'Blitz', 'Bloop', 'Lillie',  
 'Ashleigh', 'Kreggory', 'Sarge', 'Luther', 'Ivar', 'Jangle',  
 'Schnitzel', 'Panda', 'Berkeley', 'Ralphé', 'Charleson', 'Clyde',  
 'Harnold', 'Sid', 'Pippa', 'Otis', 'Carper', 'Bowie',  
 'Alexanderson', 'Suki', 'Barclay', 'Skittle', 'Ebby', 'Flávio',  
 'Smokey', 'Link', 'Jennifur', 'Ozzy', 'Bluebert', 'Stephanus',  
 'Bubbles', 'old', 'Zeus', 'Bertson', 'Nico', 'Michelangelo',  
 'Siba', 'Calbert', 'Curtis', 'Travis', 'Thumas', 'Kanu', 'Lance',  
 'Opie', 'Kane', 'Olive', 'Chuckles', 'Staniel', 'Sora', 'Beemo',  
 'Gunner', 'infuriating', 'Lacy', 'Tater', 'Olaf', 'Cecil', 'Vince',  
 'Karma', 'Billy', 'Walker', 'Rodney', 'Klevin', 'Malikai', 'Bobble',  
 'River', 'Jebberson', 'Remington', 'Farfle', 'Jiminus', 'Clarkus',  
 'Finnegus', 'Cupcake', 'Kathmandu', 'Ellie', 'Katie', 'Kara',  
 'Adele', 'Zara', 'Ambrose', 'Jimothy', 'Bode', 'Terrenth', 'Reese',  
 'Chesterson', 'Lucia', 'Bisquick', 'Ralphson', 'Socks', 'Rambo',  
 'Rudy', 'Fiji', 'Rilo', 'Bilbo', 'Coopson', 'Yoda', 'Millie',  
 'Chet', 'CROUTON', 'Daniel', 'Kaia', 'Murphy', 'Dotsy', 'Eazy',  
 'Coops', 'Fillup', 'Miley', 'Charl', 'Reagan', 'Yukon', 'CeCe',  
 'Cuddles', 'Claude', 'Jessiga', 'Carter', 'Ole', 'Pherb', 'Blipson',  
 'Reptar', 'Trevith', 'Berb', 'Bob', 'Colin', 'Brian', 'Olivier',  
 'Grady', 'Kobe', 'Freddery', 'Bodie', 'Dunkin', 'Wally', 'Tupawc',  
 'Amber', 'Edgar', 'Teddy', 'Kingsley', 'Brockly', 'Richie', 'Molly',  
 'Vinscent', 'Cedrick', 'Hazel', 'Lolo', 'Eriq', 'Phred', 'the',  
 'Oddie', 'Maxwell', 'Geoff', 'Covach', 'Durg', 'Fynn', 'Ricky',  
 'Herald', 'Lucky', 'Ferg', 'Trip', 'Clarence', 'Hamrick', 'Brad',  
 'Pubert', 'Frönq', 'Derby', 'Lizzie', 'Ember', 'Blakely', 'Opal',  
 'Marq', 'Kramer', 'Barry', 'Gordon', 'Baxter', 'Mona', 'Horace',  
 'Crimson', 'Birf', 'Hammond', 'Lorelei', 'Marty', 'Brooks',  
 'Petrick', 'Hubertson', 'Gerbald', 'Oreo', 'Bruiser', 'Perry',  
 'Bobby', 'Jeph', 'Obi', 'Tino', 'Kulet', 'Sweets', 'Lupe', 'Tiger',  
 'Jiminy', 'Griffin', 'Banjo', 'Brandy', 'Lulu', 'Darrel', 'Taco',  
 'Joey', 'Patrick', 'Kreg', 'Todo', 'Tess', 'Ulysses', 'Toffee',  
 'Apollo', 'Asher', 'Glacier', 'Chuck', 'Champ', 'Ozzie', 'Griswold',  
 'Cheesy', 'Moofasa', 'Hector', 'Goliath', 'Kawhi', 'by', 'Emmie',  
 'Penelope', 'Willie', 'Rinna', 'Mike', 'William', 'Dwight', 'Evy',  
 'officially', 'Rascal', 'Linda', 'Tug', 'Tango', 'Grizz', 'Jerome',  
 'Crumpet', 'Jessifer', 'Izzy', 'Ralph', 'Sandy', 'Humphrey',  
 'Tassy', 'Juckson', 'Chuq', 'Tyrus', 'Karl', 'Godzilla', 'Vinnie',  
 'Kenneth', 'Herm', 'Bert', 'Striker', 'Donny', 'Pepper', 'Bernie',  
 'Buddah', 'Lenny', 'Arnold', 'Zuzu', 'Mollie', 'Laela', 'Tedders',  
 'Superpup', 'Rufio', 'Jeb', 'Rodman', 'Jonah', 'Chesney', 'life',  
 'Henry', 'Bobbay', 'Mitch', 'Kaiya', 'Acro', 'Aiden', 'Obie', 'Dot',

```
'Shnuggles', 'Kendall', 'Jeffri', 'Steve', 'Mac', 'Fletcher',
'Kenzie', 'Pumpkin', 'Schnozz', 'Gustaf', 'Cheryl', 'Ed',
'Leonidas', 'Norman', 'Caryl', 'Scott', 'Taz', 'Darby', 'Jackie',
'light', 'Jazz', 'Franq', 'Pippin', 'Rolf', 'Snickers', 'Ridley',
'Cal', 'Bradley', 'Bubba', 'Tuco', 'Patch', 'Mojo', 'Batdog',
'Dylan', 'space', 'Mark', 'JD', 'Alejandro', 'Scruffers', 'Pip',
'Julius', 'Tanner', 'Sparky', 'Anthony', 'Holly', 'Jett', 'Amy',
'Sage', 'Andy', 'Mason', 'Trigger', 'Antony', 'Creg', 'Traviss',
'Gin', 'Jeffrie', 'Danny', 'Ester', 'Pluto', 'Bloo', 'Edd', 'Willy',
'Herb', 'Damon', 'Peanut', 'Nigel', 'Butters', 'Sandra', 'Fabio',
'Randall', 'Liam', 'Tommy', 'Ben', 'Raphael', 'Julio', 'Andru',
'Kloey', 'Shawwn', 'Skye', 'Kollin', 'Ronduh', 'Billl', 'Saydee',
'Dug', 'Tessa', 'Sully', 'Kirk', 'Ralf', 'Clarq', 'Jaspers',
'Samsom', 'Harrison', 'Chaz', 'Jeremy', 'Jaycob', 'Lambeau',
'Ruffles', 'Amélie', 'Bobb', 'Banditt', 'Kevon', 'Winifred', 'Hanz',
'Churlie', 'Zeek', 'Timofy', 'Maks', 'Jomathan', 'Kallie', 'Marvin',
'Spark', 'Gördón', 'Jo', 'DayZ', 'Jareld', 'Torque', 'Ron',
'Skittles', 'Cleopatra', 'Erik', 'Stu', 'Tedrick', 'Filup',
'Kial', 'Naphaniel', 'Dook', 'Hall', 'Philippe', 'Biden', 'Fwed',
'Genevieve', 'Joshwa', 'Bradlay', 'Clybe', 'Keet', 'Carll',
'Jockson', 'Josep', 'Lugan', 'Christoper'], dtype=object)
```

```
In [12]: list(df_1.name[df_1.name.isin(df_1.name[df_1.name.str.islower()])])
```

```
Out[12]: ['such',
'a',
'quite',
'quite',
'quite',
'not',
'one',
'incredibly',
'a',
'mad',
'an',
'very',
'a',
'very',
'just',
'my',
'one',
'not',
'his',
'one',
'a',
'a',
'a',
'an']
```



'very',  
'actually',  
'a',  
'just',  
'getting',  
'mad',  
'very',  
'this',  
'unacceptable',  
'all',  
'a',  
'old',  
'a',  
'infuriating',  
'a',  
'a',  
'a',  
'an',  
'a',  
'a',  
'very',  
'getting',  
'just',  
'a',  
'the',  
'the',  
'actually',  
'by',  
'a',  
'officially',  
'a',  
'the',  
'the',  
'a',  
'a',  
'a',  
'a',  
'life',  
'a',  
'one',  
'a',  
'a',  
'a',  
'light',  
'just',  
'space',  
'a',  
'the',

'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'an',  
'a',  
'the',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'a',  
'quite',  
'a',  
'an',  
'a',  
'an',  
'the',  
'the',  
'a',  
'a',  
'an',  
'a',  
'a',  
'a',  
'a']

```
In [13]: df_2.sample(50)
```

```
Out[13]:
```

	tweet_id	jpg_url
264	670797304698376195	<a href="https://pbs.twimg.com/media/CU8mlhoVAAAtes5.jpg">https://pbs.twimg.com/media/CU8mlhoVAAAtes5.jpg</a>
1913	854010172552949760	<a href="https://pbs.twimg.com/media/C9oNt91WAAAFSLs.jpg">https://pbs.twimg.com/media/C9oNt91WAAAFSLs.jpg</a>
1790	830097400375152640	<a href="https://pbs.twimg.com/media/C4UZLZLWYAA0dcs.jpg">https://pbs.twimg.com/media/C4UZLZLWYAA0dcs.jpg</a>
331	672125275208069120	<a href="https://pbs.twimg.com/media/CVPeX2dWwAEwyaR.jpg">https://pbs.twimg.com/media/CVPeX2dWwAEwyaR.jpg</a>
1763	826240494070030336	<a href="https://pbs.twimg.com/media/C3dlVMbXAAUd-Gh.jpg">https://pbs.twimg.com/media/C3dlVMbXAAUd-Gh.jpg</a>
219	670079681849372674	<a href="https://pbs.twimg.com/media/CUyZ6mVW4AI8YWZ.jpg">https://pbs.twimg.com/media/CUyZ6mVW4AI8YWZ.jpg</a>
289	671163268581498880	<a href="https://pbs.twimg.com/media/CVBzbWsWsaEynMA.jpg">https://pbs.twimg.com/media/CVBzbWsWsaEynMA.jpg</a>

885	698989035503689728	<a href="https://pbs.twimg.com/media/CbN00DaWOAARcki.jpg">https://pbs.twimg.com/media/CbN00DaWOAARcki.jpg</a>
269	670815497391357952	<a href="https://pbs.twimg.com/media/CU83IZ8W4AEIh4y.jpg">https://pbs.twimg.com/media/CU83IZ8W4AEIh4y.jpg</a>
1702	817171292965273600	<a href="https://pbs.twimg.com/media/C1cs8uAWgAEwbXc.jpg">https://pbs.twimg.com/media/C1cs8uAWgAEwbXc.jpg</a>
924	702321140488925184	<a href="https://pbs.twimg.com/media/Cb8lWafWEAA2q93.jpg">https://pbs.twimg.com/media/Cb8lWafWEAA2q93.jpg</a>
576	678675843183484930	<a href="https://pbs.twimg.com/media/CWskEqnWUAAQZW_.jpg">https://pbs.twimg.com/media/CWskEqnWUAAQZW_.jpg</a>
1216	743545585370791937	<a href="https://pbs.twimg.com/media/C1GawiUWAAAgS0w.jpg">https://pbs.twimg.com/media/C1GawiUWAAAgS0w.jpg</a>
1011	709449600415961088	<a href="https://pbs.twimg.com/media/Cdh4pgAWOAEKJ_a.jpg">https://pbs.twimg.com/media/Cdh4pgAWOAEKJ_a.jpg</a>
382	673320132811366400	<a href="https://pbs.twimg.com/media/CVgdFjNWEAAxmbq.jpg">https://pbs.twimg.com/media/CVgdFjNWEAAxmbq.jpg</a>
372	672984142909456390	<a href="https://pbs.twimg.com/media/CVbrcZyVAAA5Wpq.jpg">https://pbs.twimg.com/media/CVbrcZyVAAA5Wpq.jpg</a>
2010	878776093423087618	<a href="https://pbs.twimg.com/media/DDIKMXzWOAEibje.jpg">https://pbs.twimg.com/media/DDIKMXzWOAEibje.jpg</a>
1367	761745352076779520	<a href="https://pbs.twimg.com/media/CpJDWqhW8AAft45.jpg">https://pbs.twimg.com/media/CpJDWqhW8AAft45.jpg</a>
895	699446877801091073	<a href="https://pbs.twimg.com/media/CbTvNpoWOAEemnx.jpg">https://pbs.twimg.com/media/CbTvNpoWOAEemnx.jpg</a>
1058	714957620017307648	<a href="https://pbs.twimg.com/media/CewKKiOWwAie3pR.jpg">https://pbs.twimg.com/media/CewKKiOWwAie3pR.jpg</a>
279	671109016219725825	<a href="https://pbs.twimg.com/media/CVBCFkyU4AE2Wcr.jpg">https://pbs.twimg.com/media/CVBCFkyU4AE2Wcr.jpg</a>
994	708130923141795840	<a href="https://pbs.twimg.com/media/CdPJUWIWIAAIchl.jpg">https://pbs.twimg.com/media/CdPJUWIWIAAIchl.jpg</a>
2015	879492040517615616	<a href="https://pbs.twimg.com/media/DDSVWMvXsAEgmMK.jpg">https://pbs.twimg.com/media/DDSVWMvXsAEgmMK.jpg</a>
392	673580926094458881	<a href="https://pbs.twimg.com/media/CVkkRqOXIAEX83-.jpg">https://pbs.twimg.com/media/CVkkRqOXIAEX83-.jpg</a>
951	704871453724954624	<a href="https://pbs.twimg.com/media/Ccg02LiWEAAJHw1.jpg">https://pbs.twimg.com/media/Ccg02LiWEAAJHw1.jpg</a>
995	708149363256774660	<a href="https://pbs.twimg.com/media/CdPaEkHW8AA-Wom.jpg">https://pbs.twimg.com/media/CdPaEkHW8AA-Wom.jpg</a>
1275	750056684286914561	<a href="https://pbs.twimg.com/media/Cmfx2oNW8AAGg4H.jpg">https://pbs.twimg.com/media/Cmfx2oNW8AAGg4H.jpg</a>
748	687807801670897665	<a href="https://pbs.twimg.com/media/CYuVi9pWwAAb0GC.jpg">https://pbs.twimg.com/media/CYuVi9pWwAAb0GC.jpg</a>
323	671874878652489728	<a href="https://pbs.twimg.com/media/CVL6op1WEAAUFE7.jpg">https://pbs.twimg.com/media/CVL6op1WEAAUFE7.jpg</a>
809	692158366030913536	<a href="https://pbs.twimg.com/media/CZsKVxfWQAAxy2u.jpg">https://pbs.twimg.com/media/CZsKVxfWQAAxy2u.jpg</a>
1316	755110668769038337	<a href="https://pbs.twimg.com/ext_tw_video_thumb/75511...">https://pbs.twimg.com/ext_tw_video_thumb/75511...</a>
688	684177701129875456	<a href="https://pbs.twimg.com/media/CX6v_J0WsAE0beZ.jpg">https://pbs.twimg.com/media/CX6v_J0WsAE0beZ.jpg</a>
160	668902994700836864	<a href="https://pbs.twimg.com/media/CUhruUgUAAAa8FQ.jpg">https://pbs.twimg.com/media/CUhruUgUAAAa8FQ.jpg</a>
1610	801167903437357056	<a href="https://pbs.twimg.com/media/Cx5R8wPVEAALa9r.jpg">https://pbs.twimg.com/media/Cx5R8wPVEAALa9r.jpg</a>
598	679722016581222400	<a href="https://pbs.twimg.com/media/CW7bkW6WQAaksgB.jpg">https://pbs.twimg.com/media/CW7bkW6WQAaksgB.jpg</a>
1867	843856843873095681	<a href="https://pbs.twimg.com/media/C7X7UiOXgAA3m19.jpg">https://pbs.twimg.com/media/C7X7UiOXgAA3m19.jpg</a>
931	703079050210877440	<a href="https://pbs.twimg.com/media/CcHWqQCW8AEb0ZH.jpg">https://pbs.twimg.com/media/CcHWqQCW8AEb0ZH.jpg</a>
1378	764259802650378240	<a href="https://pbs.twimg.com/media/CpsyNtXWgAAqvs3.jpg">https://pbs.twimg.com/media/CpsyNtXWgAAqvs3.jpg</a>
1658	810984652412424192	<a href="https://pbs.twimg.com/media/COEypZbXAAaceSc.jpg">https://pbs.twimg.com/media/COEypZbXAAaceSc.jpg</a>
638	681281657291280384	<a href="https://pbs.twimg.com/media/CXRmDfWWMAADCdc.jpg">https://pbs.twimg.com/media/CXRmDfWWMAADCdc.jpg</a>
1629	805207613751304193	<a href="https://pbs.twimg.com/media/CyysDQlVIAAYgrl.jpg">https://pbs.twimg.com/media/CyysDQlVIAAYgrl.jpg</a>
1006	709158332880297985	<a href="https://pbs.twimg.com/media/CddvvSwWoAU0bQw.jpg">https://pbs.twimg.com/media/CddvvSwWoAU0bQw.jpg</a>
1970	868880397819494401	<a href="https://pbs.twimg.com/media/DA7iHL5U0AA10Qo.jpg">https://pbs.twimg.com/media/DA7iHL5U0AA10Qo.jpg</a>
1886	847962785489326080	<a href="https://pbs.twimg.com/media/C8SRpHNUIAARB3j.jpg">https://pbs.twimg.com/media/C8SRpHNUIAARB3j.jpg</a>
636	681242418453299201	<a href="https://pbs.twimg.com/media/CXRCXesVAAArSXt.jpg">https://pbs.twimg.com/media/CXRCXesVAAArSXt.jpg</a>
1053	714214115368108032	<a href="https://pbs.twimg.com/media/Cell8ikWIAACCJ-.jpg">https://pbs.twimg.com/media/Cell8ikWIAACCJ-.jpg</a>
507	676089483918516224	<a href="https://pbs.twimg.com/media/CWHzzFGXIAA0Y_H.jpg">https://pbs.twimg.com/media/CWHzzFGXIAA0Y_H.jpg</a>
812	692530551048294401	<a href="https://pbs.twimg.com/media/CZxc3G7WEAAM4Mv.jpg">https://pbs.twimg.com/media/CZxc3G7WEAAM4Mv.jpg</a>
1996	874680097055178752	<a href="https://pbs.twimg.com/media/DCN85nGUwAAzG_q.jpg">https://pbs.twimg.com/media/DCN85nGUwAAzG_q.jpg</a>
1445	775842724423557120	<a href="https://pbs.twimg.com/media/CsRY1jAWYAU0x55.jpg">https://pbs.twimg.com/media/CsRY1jAWYAU0x55.jpg</a>

	img_num	p1	p1_conf	p1_dog	\
264	1	Pembroke	0.472197	True	
1913	1	English_springer	0.354733	True	
1790	4	toy_poodle	0.442713	True	

331	1	tennis_ball	0.999834	False
1763	1	French_bulldog	0.903048	True
219	1	mud_turtle	0.157477	False
289	1	African_hunting_dog	0.733025	False
885	1	Norfolk_terrier	0.246340	True
269	1	American_Staffordshire_terrier	0.919714	True
1702	1	golden_retriever	0.295483	True
924	3	West_Highland_white_terrier	0.769159	True
576	1	maze	0.339850	False
1216	2	rapeseed	0.876875	False
1011	2	Maltese_dog	0.780187	True
382	3	Samoyed	0.978833	True
372	1	wombat	0.738780	False
2010	2	Italian_greyhound	0.734684	True
1367	1	paddle	0.393118	False
895	3	Pembroke	0.969400	True
1058	1	Great_Pyrenees	0.251516	True
279	1	basenji	0.855959	True
994	1	French_bulldog	0.710354	True
2015	1	German_short-haired_pointer	0.479896	True
392	1	beagle	0.985062	True
951	1	Norfolk_terrier	0.689504	True
995	1	Cardigan	0.350993	True
1275	1	Saluki	0.484428	True
748	1	Staffordshire_bullterrier	0.151113	True
323	1	china_cabinet	0.996031	False
809	1	pug	0.956565	True
1316	1	Labrador_retriever	0.708974	True
688	1	chow	0.334783	True
160	1	Brittany_spaniel	0.828425	True
1610	1	cocker_spaniel	0.740220	True
598	1	boxer	0.459604	True
1867	1	Labrador_retriever	0.922540	True
931	2	Pembroke	0.778503	True
1378	1	German_shepherd	0.973677	True
1658	1	golden_retriever	0.871342	True
638	1	Saint_Bernard	0.998830	True
1629	1	Pembroke	0.244705	True
1006	1	Siberian_husky	0.212957	True
1970	1	laptop	0.153718	False
1886	1	sea_lion	0.882654	False
636	1	motor_scooter	0.255934	False
1053	1	pug	0.533967	True
507	1	bull_mastiff	0.743808	True
812	1	Siberian_husky	0.486428	True
1996	1	Labrador_retriever	0.836052	True
1445	2	chow	0.520022	True

	p2	p2_conf	p2_dog \
264	beagle	0.090938	True
1913	collie	0.177538	True
1790	Pomeranian	0.142073	True
331	golden_retriever	0.000087	True
1763	pug	0.096242	True
219	terrapin	0.131846	False
289	plow	0.119377	False
885	Irish_terrier	0.243349	True
269	Staffordshire_bullterrier	0.073430	True
1702	Irish_setter	0.144431	True
924	Scotch_terrier	0.064369	True
576	streetcar	0.099688	False
1216	standard_poodle	0.060584	True
1011	Dandie_Dinmont	0.074429	True
382	Pomeranian	0.012763	True
372	beaver	0.133368	False
2010	whippet	0.150487	True
1367	canoe	0.178088	False
895	Cardigan	0.026059	True
1058	Samoyed	0.139346	True
279	beagle	0.036723	True
994	Chihuahua	0.262302	True
2015	vizsla	0.124353	True
392	basset	0.006418	True
951	soft-coated_wheaten_terrier	0.101480	True
995	basset	0.164555	True
1275	borzoi	0.263550	True
748	boxer	0.135697	True
323	entertainment_center	0.001986	False
809	swing	0.018907	False
1316	golden_retriever	0.114314	True
688	German_shepherd	0.162647	True
160	Ibizan_hound	0.043082	True
1610	Dandie_Dinmont	0.061604	True
598	Boston_bull	0.197913	True
1867	golden_retriever	0.074358	True
931	Shetland_sheepdog	0.093834	True
1378	malinois	0.025950	True
1658	Tibetan_mastiff	0.036708	True
638	Pekinese	0.000391	True
1629	Rhodesian_ridgeback	0.180461	True
1006	Eskimo_dog	0.178887	True
1970	French_bulldog	0.099984	True
1886	mink	0.066880	False
636	rifle	0.145202	False
1053	bloodhound	0.164826	True
507	boxer	0.106697	True

812	Eskimo_dog	0.448518	True
1996	Staffordshire_bullterrier	0.047069	True
1445	bath_towel	0.028775	False

		p3	p3_conf	p3_dog
264	German_shepherd	0.064366		True
1913	Border_collie	0.131706		True
1790	Pekinese	0.125745		True
331	racket	0.000053		False
1763	Boston_bull	0.000234		True
219	box_turtle	0.060678		False
289	Scottish_deerhound	0.026983		True
885	golden_retriever	0.085871		True
269	bull_mastiff	0.000906		True
1702	Chesapeake_Bay_retriever	0.077879		True
924	Old_English_sheepdog	0.043763		True
576	sundial	0.084808		False
1216	Great_Pyrenees	0.033006		True
1011	Norfolk_terrier	0.033776		True
382	Eskimo_dog	0.001853		True
372	wallaby	0.032370		False
2010	Ibizan_hound	0.039725		True
1367	lakeside	0.099713		False
895	Chihuahua	0.003505		True
1058	kuvasz	0.129005		True
279	toy_terrier	0.029258		True
994	Cardigan	0.006904		True
2015	bath_towel	0.073320		False
392	Walker_hound	0.003533		True
951	Norwich_terrier	0.055779		True
995	toy_terrier	0.080484		True
1275	Labrador_retriever	0.077004		True
748	American_Staffordshire_terrier	0.086591		True
323	bookcase	0.001652		False
809	toy_poodle	0.013544		True
1316	Great_Pyrenees	0.065813		True
688	golden_retriever	0.138612		True
160	Blenheim_spaniel	0.028004		True
1610	English_setter	0.041331		True
598	French_bulldog	0.087023		True
1867	Great_Pyrenees	0.002325		True
931	Cardigan	0.060296		True
1378	kelpie	0.000192		True
1658	Labrador_retriever	0.025823		True
638	Great_Pyrenees	0.000224		True
1629	Cardigan	0.094664		True
1006	Labrador_retriever	0.174218		True
1970	printer	0.077130		False

1886	otter	0.025679	False
636	assault_rifle	0.097000	False
1053	German_shepherd	0.046524	True
507	American_Staffordshire_terrier	0.042335	True
812	white_wolf	0.041506	False
1996	beagle	0.036007	True
1445	French_bulldog	0.025990	True

In [14]: df\_2.shape

Out[14]: (2075, 12)

In [15]: df\_2.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 [16]: df\_2.describe()

Out[16]:	tweet_id	img_num	p1_conf	p2_conf	p3_conf
count	2.075000e+03	2075.000000	2075.000000	2.075000e+03	2.075000e+03
mean	7.384514e+17	1.203855	0.594548	1.345886e-01	6.032417e-02
std	6.785203e+16	0.561875	0.271174	1.006657e-01	5.090593e-02
min	6.660209e+17	1.000000	0.044333	1.011300e-08	1.740170e-10
25%	6.764835e+17	1.000000	0.364412	5.388625e-02	1.622240e-02
50%	7.119988e+17	1.000000	0.588230	1.181810e-01	4.944380e-02
75%	7.932034e+17	1.000000	0.843855	1.955655e-01	9.180755e-02
max	8.924206e+17	4.000000	1.000000	4.880140e-01	2.734190e-01

In [17]: df\_2['img\_num'].value\_counts()

Out[17]:	1	1780
	2	198

```
3      66
4      31
Name: img_num, dtype: int64
```

```
In [18]: df_3.sample(20)
```

```
Out[18]:
```

	favorite_count	retweet_count	tweet_id
1018	10520	4432	746757706116112384
213	25944	7855	851464819735769094
1633	4225	1563	684222868335505415
1821	2025	621	676533798876651520
103	35085	6649	872261713294495745
1465	2203	569	694329668942569472
840	5650	1565	766793450729734144
1794	1362	471	677301033169788928
352	12819	2786	831309418084069378
1730	39726	15839	679828447187857408
1345	8804	3999	704364645503647744
2089	872	431	670783437142401025
621	9469	2703	796080075804475393
235	22036	4800	847251039262605312
1795	2164	790	677269281705472000
907	5022	1391	757725642876129280
869	4578	1336	761599872357261312
689	12120	3300	787397959788929025
2152	3769	1963	669625907762618368
651	15128	4998	792394556390137856

```
In [19]: df_3.isnull().sum()
```

```
Out[19]: favorite_count    0
         retweet_count    0
         tweet_id        0
         dtype: int64
```

```
In [20]: df_3.info()
```

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

```
In [21]: df_3.describe()
```



```
Out[21]:
```

	favorite_count	retweet_count	tweet_id
count	2354.000000	2354.000000	2.354000e+03
mean	8080.968564	3164.797366	7.426978e+17
std	11814.771334	5284.770364	6.852812e+16
min	0.000000	0.000000	6.660209e+17
25%	1415.000000	624.500000	6.783975e+17
50%	3603.500000	1473.500000	7.194596e+17
75%	10122.250000	3652.000000	7.993058e+17
max	132810.000000	79515.000000	8.924206e+17

## 2 After Visaul and programatic assessment .

### 2.0.1 1.2.1 Quality

'df\_1' dataset :

- \* [1] (in\_reply\_to\_status\_id ,in\_reply\_to\_user\_id) not null values == validity issue.
- \* [2] (retweeted\_status\_id, retweeted\_status\_user\_id, retweeted\_status\_timestamp) not null values == validity issue
- \* [3] timestamp column isn't date time.
- \* [4] extract source of tweet from source column and categorize it. \* [5] text column contain urls. \* [6] expanded\_urls have NaN values. \* [7] some expanded\_urls contain duplicated urls in the same value. \* [8] some rating\_denominator value == accuracy issue. \* [9] some rating\_numerator value not == validity issue. \* [10] doggo,floofer,pupper,puppo columns has 'none' for missing values.
- \* [11] some name column values == validity issue.

'df\_2' dataset :

- \* [12] img\_num column == consistency . \* [13] df\_2 should contain only the highest prediction confident of breed and the breed name instead of 3 predictions.

### 2.0.2 1.2.2 Tidness

- [1] 'df\_1' table : doggo,floofer,pupper,puppo are variable not columns.
- [2] All should be in one dataset.

## 2.1 1.3 Clean

```
In [22]: df1_clean = df_1.copy()
         df2_clean = df_2.copy()
         df3_clean = df_3.copy()
```

### 2.1.1 1.3.1 Quality

Define

- 'df\_1' dataset :(in\_reply\_to\_status\_id ,in\_reply\_to\_user\_id) not\_null values == validity issue.

## Code

```
In [23]: ## null values are the required rows so I used is_null method to drop not_null rows.
```

```
df1_clean= df1_clean[df1_clean['in_reply_to_status_id'].isnull()]
df1_clean= df1_clean[df1_clean['in_reply_to_user_id'].isnull()]
```

```
df1_clean
```

```
Out[23]:
```

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

2334	666293911632134144	NaN	NaN
2335	666287406224695296	NaN	NaN
2336	666273097616637952	NaN	NaN
2337	666268910803644416	NaN	NaN
2338	666104133288665088	NaN	NaN
2339	666102155909144576	NaN	NaN
2340	666099513787052032	NaN	NaN
2341	666094000022159362	NaN	NaN
2342	666082916733198337	NaN	NaN
2343	666073100786774016	NaN	NaN
2344	666071193221509120	NaN	NaN
2345	666063827256086533	NaN	NaN
2346	666058600524156928	NaN	NaN
2347	666057090499244032	NaN	NaN
2348	666055525042405380	NaN	NaN
2349	666051853826850816	NaN	NaN
2350	666050758794694657	NaN	NaN
2351	666049248165822465	NaN	NaN
2352	666044226329800704	NaN	NaN
2353	666033412701032449	NaN	NaN
2354	666029285002620928	NaN	NaN
2355	666020888022790149	NaN	NaN

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

24 2017-07-18 16:08:03 +0000  
 25 2017-07-18 00:07:08 +0000  
 26 2017-07-17 16:17:36 +0000  
 27 2017-07-16 23:58:41 +0000  
 28 2017-07-16 20:14:00 +0000  
 29 2017-07-15 23:25:31 +0000  
 ... ..  
 2326 2015-11-17 00:24:19 +0000  
 2327 2015-11-17 00:06:54 +0000  
 2328 2015-11-16 23:23:41 +0000  
 2329 2015-11-16 21:54:18 +0000  
 2330 2015-11-16 21:10:36 +0000  
 2331 2015-11-16 20:32:58 +0000  
 2332 2015-11-16 20:01:42 +0000  
 2333 2015-11-16 19:31:45 +0000  
 2334 2015-11-16 16:37:02 +0000  
 2335 2015-11-16 16:11:11 +0000  
 2336 2015-11-16 15:14:19 +0000  
 2337 2015-11-16 14:57:41 +0000  
 2338 2015-11-16 04:02:55 +0000  
 2339 2015-11-16 03:55:04 +0000  
 2340 2015-11-16 03:44:34 +0000  
 2341 2015-11-16 03:22:39 +0000  
 2342 2015-11-16 02:38:37 +0000  
 2343 2015-11-16 01:59:36 +0000  
 2344 2015-11-16 01:52:02 +0000  
 2345 2015-11-16 01:22:45 +0000  
 2346 2015-11-16 01:01:59 +0000  
 2347 2015-11-16 00:55:59 +0000  
 2348 2015-11-16 00:49:46 +0000  
 2349 2015-11-16 00:35:11 +0000  
 2350 2015-11-16 00:30:50 +0000  
 2351 2015-11-16 00:24:50 +0000  
 2352 2015-11-16 00:04:52 +0000  
 2353 2015-11-15 23:21:54 +0000  
 2354 2015-11-15 23:05:30 +0000  
 2355 2015-11-15 22:32:08 +0000

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

[illegible]

2352 <a href="http://twitter.com/download/iphone" r...  
 2353 <a href="http://twitter.com/download/iphone" r...  
 2354 <a href="http://twitter.com/download/iphone" r...  
 2355 <a href="http://twitter.com/download/iphone" r...

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

2337	Very concerned about fellow dog trapped in com...	NaN
2338	Not familiar with this breed. No tail (weird)...	NaN
2339	Oh my. Here you are seeing an Adobe Setter giv...	NaN
2340	Can stand on stump for what seems like a while...	NaN
2341	This appears to be a Mongolian Presbyterian mi...	NaN
2342	Here we have a well-established sunblockerspan...	NaN
2343	Let's hope this flight isn't Malaysian (lol). ...	NaN
2344	Here we have a northern speckled Rhododendron...	NaN
2345	This is the happiest dog you will ever see. Ve...	NaN
2346	Here is the Rand Paul of retrievers folks! He'...	NaN
2347	My oh my. This is a rare blond Canadian terrie...	NaN
2348	Here is a Siberian heavily armored polar bear ...	NaN
2349	This is an odd dog. Hard on the outside but lo...	NaN
2350	This is a truly beautiful English Wilson Staff...	NaN
2351	Here we have a 1949 1st generation vulpix. Enj...	NaN
2352	This is a purebred Piers Morgan. Loves to Netf...	NaN
2353	Here is a very happy pup. Big fan of well-main...	NaN
2354	This is a western brown Mitsubishi terrier. Up...	NaN
2355	Here we have a Japanese Irish Setter. Lost eye...	NaN

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

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

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



12	<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://...">https://...</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

	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

[2278 rows x 17 columns]

## Test

In [24]: df1\_clean.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2278 entries, 0 to 2355
Data columns (total 17 columns):
tweet_id                2278 non-null int64
in_reply_to_status_id    0 non-null float64
in_reply_to_user_id      0 non-null float64
timestamp                2278 non-null object
source                  2278 non-null object
text                    2278 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            2274 non-null object
rating_numerator         2278 non-null int64
rating_denominator       2278 non-null int64
name                     2278 non-null object
doggo                    2278 non-null object
floofer                  2278 non-null object
pupper                   2278 non-null object
puppo                    2278 non-null object
dtypes: float64(4), int64(3), object(10)
memory usage: 320.3+ KB
```

## Define

- 'df\_1' dataset : (retweeted\_status\_id, retweeted\_status\_user\_id, retweeted\_status\_timestamp) not\_null values == validity issue.

## Code

In [25]: *## null values are the required rows so I used is\_null method to drop not\_null rows.*

```
df1_clean= df1_clean[df1_clean['retweeted_status_id'].isnull()]
df1_clean= df1_clean[df1_clean['retweeted_status_user_id'].isnull()]
df1_clean= df1_clean[df1_clean['retweeted_status_timestamp'].isnull()]
```

df1\_clean

```
Out[25]:
```

	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
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
31	886258384151887873	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

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  
 31 2017-07-15 16:17:19 +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...

[illegible]

2347 <a href="http://twitter.com/download/iphone" r...  
 2348 <a href="http://twitter.com/download/iphone" r...  
 2349 <a href="http://twitter.com/download/iphone" r...  
 2350 <a href="http://twitter.com/download/iphone" r...  
 2351 <a href="http://twitter.com/download/iphone" r...  
 2352 <a href="http://twitter.com/download/iphone" r...  
 2353 <a href="http://twitter.com/download/iphone" r...  
 2354 <a href="http://twitter.com/download/iphone" r...  
 2355 <a href="http://twitter.com/download/iphone" r...

	text	retweeted_status_id \
0	This is Phineas. He's a mystical boy. Only eve...	NaN
1	This is Tilly. She's just checking pup on you...	NaN
2	This is Archie. He is a rare Norwegian Pouncin...	NaN
3	This is Darla. She commenced a snooze mid meal...	NaN
4	This is Franklin. He would like you to stop ca...	NaN
5	Here we have a majestic great white breaching ...	NaN
6	Meet Jax. He enjoys ice cream so much he gets ...	NaN
7	When you watch your owner call another dog a g...	NaN
8	This is Zoey. She doesn't want to be one of th...	NaN
9	This is Cassie. She is a college pup. Studying...	NaN
10	This is Koda. He is a South Australian decksha...	NaN
11	This is Bruno. He is a service shark. Only get...	NaN
12	Here's a puppo that seems to be on the fence a...	NaN
13	This is Ted. He does his best. Sometimes that'...	NaN
14	This is Stuart. He's sporting his favorite fan...	NaN
15	This is Oliver. You're witnessing one of his m...	NaN
16	This is Jim. He found a fren. Taught him how t...	NaN
17	This is Zeke. He has a new stick. Very proud o...	NaN
18	This is Ralphus. He's powering up. Attempting ...	NaN
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
31	This is Waffles. His doggles are pupside down...	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	
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
31	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
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://...">https://...</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
31	<a href="https://twitter.com/dog_rates/status/886258384...">https://twitter.com/dog_rates/status/886258384...</a>	13
...	...	...
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
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
31	10	Waffles	None	None	None	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

[2097 rows x 17 columns]

## Test

In [26]: df1\_clean.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 17 columns):
tweet_id                2097 non-null int64
in_reply_to_status_id    0 non-null float64
in_reply_to_user_id      0 non-null float64
timestamp                2097 non-null object
source                  2097 non-null object
text                    2097 non-null object
retweeted_status_id      0 non-null float64
retweeted_status_user_id 0 non-null float64
retweeted_status_timestamp 0 non-null object
expanded_urls            2094 non-null object
rating_numerator          2097 non-null int64
rating_denominator        2097 non-null int64
name                     2097 non-null object
doggo                    2097 non-null object
floofer                  2097 non-null object
pupper                   2097 non-null object
puppo                     2097 non-null object
```

```
dtypes: float64(4), int64(3), object(10)
memory usage: 294.9+ KB
```

## Define

- columns ['retweeted\_status\_id', 'retweeted\_status\_user\_id', 'retweeted\_status\_timestamp', 'in\_reply\_to\_status\_id', 'in\_reply\_to\_user\_id'] not required

## Code

```
In [27]: ## drop columns ['retweeted_status_id', 'retweeted_status_user_id', 'retweeted_status_t
df1_clean = df1_clean.drop(columns=['retweeted_status_id', 'retweeted_status_user_id',
                                     'in_reply_to_status_id', 'in_reply_to_user_id'], a
```

```
df1_clean
```

```
Out[27]:
```

	tweet_id	timestamp \
0	892420643555336193	2017-08-01 16:23:56 +0000
1	892177421306343426	2017-08-01 00:17:27 +0000
2	891815181378084864	2017-07-31 00:18:03 +0000
3	891689557279858688	2017-07-30 15:58:51 +0000
4	891327558926688256	2017-07-29 16:00:24 +0000
5	891087950875897856	2017-07-29 00:08:17 +0000
6	890971913173991426	2017-07-28 16:27:12 +0000
7	890729181411237888	2017-07-28 00:22:40 +0000
8	890609185150312448	2017-07-27 16:25:51 +0000
9	890240255349198849	2017-07-26 15:59:51 +0000
10	890006608113172480	2017-07-26 00:31:25 +0000
11	889880896479866881	2017-07-25 16:11:53 +0000
12	889665388333682689	2017-07-25 01:55:32 +0000
13	889638837579907072	2017-07-25 00:10:02 +0000
14	889531135344209921	2017-07-24 17:02:04 +0000
15	889278841981685760	2017-07-24 00:19:32 +0000
16	888917238123831296	2017-07-23 00:22:39 +0000
17	888804989199671297	2017-07-22 16:56:37 +0000
18	888554962724278272	2017-07-22 00:23:06 +0000
20	888078434458587136	2017-07-20 16:49:33 +0000
21	887705289381826560	2017-07-19 16:06:48 +0000
22	887517139158093824	2017-07-19 03:39:09 +0000
23	887473957103951883	2017-07-19 00:47:34 +0000
24	887343217045368832	2017-07-18 16:08:03 +0000
25	887101392804085760	2017-07-18 00:07:08 +0000
26	886983233522544640	2017-07-17 16:17:36 +0000
27	886736880519319552	2017-07-16 23:58:41 +0000
28	886680336477933568	2017-07-16 20:14:00 +0000
29	886366144734445568	2017-07-15 23:25:31 +0000
31	886258384151887873	2017-07-15 16:17:19 +0000

```

...
2326 666411507551481857 2015-11-17 00:24:19 +0000
2327 666407126856765440 2015-11-17 00:06:54 +0000
2328 666396247373291520 2015-11-16 23:23:41 +0000
2329 666373753744588802 2015-11-16 21:54:18 +0000
2330 666362758909284353 2015-11-16 21:10:36 +0000
2331 666353288456101888 2015-11-16 20:32:58 +0000
2332 666345417576210432 2015-11-16 20:01:42 +0000
2333 666337882303524864 2015-11-16 19:31:45 +0000
2334 666293911632134144 2015-11-16 16:37:02 +0000
2335 666287406224695296 2015-11-16 16:11:11 +0000
2336 666273097616637952 2015-11-16 15:14:19 +0000
2337 666268910803644416 2015-11-16 14:57:41 +0000
2338 666104133288665088 2015-11-16 04:02:55 +0000
2339 666102155909144576 2015-11-16 03:55:04 +0000
2340 666099513787052032 2015-11-16 03:44:34 +0000
2341 666094000022159362 2015-11-16 03:22:39 +0000
2342 666082916733198337 2015-11-16 02:38:37 +0000
2343 666073100786774016 2015-11-16 01:59:36 +0000
2344 666071193221509120 2015-11-16 01:52:02 +0000
2345 666063827256086533 2015-11-16 01:22:45 +0000
2346 666058600524156928 2015-11-16 01:01:59 +0000
2347 666057090499244032 2015-11-16 00:55:59 +0000
2348 666055525042405380 2015-11-16 00:49:46 +0000
2349 666051853826850816 2015-11-16 00:35:11 +0000
2350 666050758794694657 2015-11-16 00:30:50 +0000
2351 666049248165822465 2015-11-16 00:24:50 +0000
2352 666044226329800704 2015-11-16 00:04:52 +0000
2353 666033412701032449 2015-11-15 23:21:54 +0000
2354 666029285002620928 2015-11-15 23:05:30 +0000
2355 666020888022790149 2015-11-15 22:32:08 +0000

```

```

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

```

[illegible]

text \



0 This is Phineas. He's a mystical boy. Only eve...  
 1 This is Tilly. She's just checking pup on you...  
 2 This is Archie. He is a rare Norwegian Pouncin...  
 3 This is Darla. She commenced a snooze mid meal...  
 4 This is Franklin. He would like you to stop ca...  
 5 Here we have a majestic great white breaching ...  
 6 Meet Jax. He enjoys ice cream so much he gets ...  
 7 When you watch your owner call another dog a g...  
 8 This is Zoey. She doesn't want to be one of th...  
 9 This is Cassie. She is a college pup. Studying...  
 10 This is Koda. He is a South Australian decksha...  
 11 This is Bruno. He is a service shark. Only get...  
 12 Here's a puppo that seems to be on the fence a...  
 13 This is Ted. He does his best. Sometimes that'...  
 14 This is Stuart. He's sporting his favorite fan...  
 15 This is Oliver. You're witnessing one of his m...  
 16 This is Jim. He found a fren. Taught him how t...  
 17 This is Zeke. He has a new stick. Very proud o...  
 18 This is Ralphus. He's powering up. Attempting ...  
 20 This is Gerald. He was just told he didn't get...  
 21 This is Jeffrey. He has a monopoly on the pool...  
 22 I've yet to rate a Venezuelan Hover Wiener. Th...  
 23 This is Canela. She attempted some fancy porch...  
 24 You may not have known you needed to see this ...  
 25 This... is a Jubilant Antarctic House Bear. We...  
 26 This is Maya. She's very shy. Rarely leaves he...  
 27 This is Mingus. He's a wonderful father to his...  
 28 This is Derek. He's late for a dog meeting. 13...  
 29 This is Roscoe. Another pupper fallen victim t...  
 31 This is Waffles. His doggles are pupside down...  
 ...  
 2326 This is quite the dog. Gets really excited whe...  
 2327 This is a southern Vesuvius bumblegruff. Can d...  
 2328 Oh goodness. A super rare northeast Qdoba kang...  
 2329 Those are sunglasses and a jean jacket. 11/10 ...  
 2330 Unique dog here. Very small. Lives in containe...  
 2331 Here we have a mixed Asiago from the Galápagos...  
 2332 Look at this jokester thinking seat belt laws ...  
 2333 This is an extremely rare horned Parthenon. No...  
 2334 This is a funny dog. Weird toes. Won't come do...  
 2335 This is an Albanian 3 1/2 legged Episcopalian...  
 2336 Can take selfies 11/10 <https://t.co/ws2AMaWpW>  
 2337 Very concerned about fellow dog trapped in com...  
 2338 Not familiar with this breed. No tail (weird)...  
 2339 Oh my. Here you are seeing an Adobe Setter giv...  
 2340 Can stand on stump for what seems like a while...  
 2341 This appears to be a Mongolian Presbyterian mi...  
 2342 Here we have a well-established sunblockerspan...

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

	expanded_urls	rating_numerator \
0	https://twitter.com/dog_rates/status/892420643...	13
1	https://twitter.com/dog_rates/status/892177421...	13
2	https://twitter.com/dog_rates/status/891815181...	12
3	https://twitter.com/dog_rates/status/891689557...	13
4	https://twitter.com/dog_rates/status/891327558...	12
5	https://twitter.com/dog_rates/status/891087950...	13
6	https://gofundme.com/ydvmve-surgery-for-jax,ht...	13
7	https://twitter.com/dog_rates/status/890729181...	13
8	https://twitter.com/dog_rates/status/890609185...	13
9	https://twitter.com/dog_rates/status/890240255...	14
10	https://twitter.com/dog_rates/status/890006608...	13
11	https://twitter.com/dog_rates/status/889880896...	13
12	https://twitter.com/dog_rates/status/889665388...	13
13	https://twitter.com/dog_rates/status/889638837...	12
14	https://twitter.com/dog_rates/status/889531135...	13
15	https://twitter.com/dog_rates/status/889278841...	13
16	https://twitter.com/dog_rates/status/888917238...	12
17	https://twitter.com/dog_rates/status/888804989...	13
18	https://twitter.com/dog_rates/status/888554962...	13
20	https://twitter.com/dog_rates/status/888078434...	12
21	https://twitter.com/dog_rates/status/887705289...	13
22	https://twitter.com/dog_rates/status/887517139...	14
23	https://twitter.com/dog_rates/status/887473957...	13
24	https://twitter.com/dog_rates/status/887343217...	13
25	https://twitter.com/dog_rates/status/887101392...	12
26	https://twitter.com/dog_rates/status/886983233...	13
27	https://www.gofundme.com/mingusneedsus,https://...	13
28	https://twitter.com/dog_rates/status/886680336...	13
29	https://twitter.com/dog_rates/status/886366144...	12
31	https://twitter.com/dog_rates/status/886258384...	13
...	...	...
2326	https://twitter.com/dog_rates/status/666411507...	2
2327	https://twitter.com/dog_rates/status/666407126...	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
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
31	10	Waffles	None	None	None	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

[2097 rows x 12 columns]

Test

```
In [28]: df1_clean.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 12 columns):
tweet_id          2097 non-null int64
timestamp         2097 non-null object
source            2097 non-null object
text              2097 non-null object
expanded_urls     2094 non-null object
rating_numerator  2097 non-null int64
rating_denominator 2097 non-null int64
name              2097 non-null object
doggo             2097 non-null object
floofer          2097 non-null object
pupper           2097 non-null object
puppo            2097 non-null object
dtypes: int64(3), object(9)
memory usage: 213.0+ KB
```

## Define

- 'df\_1' dataset : timestamp column is string instead of date time.

## Code

```
In [29]: ## use to_datetime method
```

```
df1_clean['timestamp'] = pd.to_datetime(df1_clean['timestamp'])
df1_clean
```

```
Out[29]:
```

	tweet_id	timestamp	\
0	892420643555336193	2017-08-01 16:23:56	
1	892177421306343426	2017-08-01 00:17:27	
2	891815181378084864	2017-07-31 00:18:03	
3	891689557279858688	2017-07-30 15:58:51	
4	891327558926688256	2017-07-29 16:00:24	
5	891087950875897856	2017-07-29 00:08:17	
6	890971913173991426	2017-07-28 16:27:12	
7	890729181411237888	2017-07-28 00:22:40	
8	890609185150312448	2017-07-27 16:25:51	
9	890240255349198849	2017-07-26 15:59:51	
10	890006608113172480	2017-07-26 00:31:25	
11	889880896479866881	2017-07-25 16:11:53	
12	889665388333682689	2017-07-25 01:55:32	
13	889638837579907072	2017-07-25 00:10:02	
14	889531135344209921	2017-07-24 17:02:04	

15	889278841981685760	2017-07-24	00:19:32
16	888917238123831296	2017-07-23	00:22:39
17	888804989199671297	2017-07-22	16:56:37
18	888554962724278272	2017-07-22	00:23:06
20	888078434458587136	2017-07-20	16:49:33
21	887705289381826560	2017-07-19	16:06:48
22	887517139158093824	2017-07-19	03:39:09
23	887473957103951883	2017-07-19	00:47:34
24	887343217045368832	2017-07-18	16:08:03
25	887101392804085760	2017-07-18	00:07:08
26	886983233522544640	2017-07-17	16:17:36
27	886736880519319552	2017-07-16	23:58:41
28	886680336477933568	2017-07-16	20:14:00
29	886366144734445568	2017-07-15	23:25:31
31	886258384151887873	2017-07-15	16:17:19
...	...	...	...
2326	666411507551481857	2015-11-17	00:24:19
2327	666407126856765440	2015-11-17	00:06:54
2328	666396247373291520	2015-11-16	23:23:41
2329	666373753744588802	2015-11-16	21:54:18
2330	666362758909284353	2015-11-16	21:10:36
2331	666353288456101888	2015-11-16	20:32:58
2332	666345417576210432	2015-11-16	20:01:42
2333	666337882303524864	2015-11-16	19:31:45
2334	666293911632134144	2015-11-16	16:37:02
2335	666287406224695296	2015-11-16	16:11:11
2336	666273097616637952	2015-11-16	15:14:19
2337	666268910803644416	2015-11-16	14:57:41
2338	666104133288665088	2015-11-16	04:02:55
2339	666102155909144576	2015-11-16	03:55:04
2340	666099513787052032	2015-11-16	03:44:34
2341	666094000022159362	2015-11-16	03:22:39
2342	666082916733198337	2015-11-16	02:38:37
2343	666073100786774016	2015-11-16	01:59:36
2344	666071193221509120	2015-11-16	01:52:02
2345	666063827256086533	2015-11-16	01:22:45
2346	666058600524156928	2015-11-16	01:01:59
2347	666057090499244032	2015-11-16	00:55:59
2348	666055525042405380	2015-11-16	00:49:46
2349	666051853826850816	2015-11-16	00:35:11
2350	666050758794694657	2015-11-16	00:30:50
2351	666049248165822465	2015-11-16	00:24:50
2352	666044226329800704	2015-11-16	00:04:52
2353	666033412701032449	2015-11-15	23:21:54
2354	666029285002620928	2015-11-15	23:05:30
2355	666020888022790149	2015-11-15	22:32:08

source \

[illegible]

2343 <a href="http://twitter.com/download/iphone" r...  
 2344 <a href="http://twitter.com/download/iphone" r...  
 2345 <a href="http://twitter.com/download/iphone" r...  
 2346 <a href="http://twitter.com/download/iphone" r...  
 2347 <a href="http://twitter.com/download/iphone" r...  
 2348 <a href="http://twitter.com/download/iphone" r...  
 2349 <a href="http://twitter.com/download/iphone" r...  
 2350 <a href="http://twitter.com/download/iphone" r...  
 2351 <a href="http://twitter.com/download/iphone" r...  
 2352 <a href="http://twitter.com/download/iphone" r...  
 2353 <a href="http://twitter.com/download/iphone" r...  
 2354 <a href="http://twitter.com/download/iphone" r...  
 2355 <a href="http://twitter.com/download/iphone" r...

text \

0 This is Phineas. He's a mystical boy. Only eve...  
 1 This is Tilly. She's just checking pup on you...  
 2 This is Archie. He is a rare Norwegian Pouncin...  
 3 This is Darla. She commenced a snooze mid meal...  
 4 This is Franklin. He would like you to stop ca...  
 5 Here we have a majestic great white breaching ...  
 6 Meet Jax. He enjoys ice cream so much he gets ...  
 7 When you watch your owner call another dog a g...  
 8 This is Zoey. She doesn't want to be one of th...  
 9 This is Cassie. She is a college pup. Studying...  
 10 This is Koda. He is a South Australian decksha...  
 11 This is Bruno. He is a service shark. Only get...  
 12 Here's a puppo that seems to be on the fence a...  
 13 This is Ted. He does his best. Sometimes that'...  
 14 This is Stuart. He's sporting his favorite fan...  
 15 This is Oliver. You're witnessing one of his m...  
 16 This is Jim. He found a fren. Taught him how t...  
 17 This is Zeke. He has a new stick. Very proud o...  
 18 This is Ralphus. He's powering up. Attempting ...  
 20 This is Gerald. He was just told he didn't get...  
 21 This is Jeffrey. He has a monopoly on the pool...  
 22 I've yet to rate a Venezuelan Hover Wiener. Th...  
 23 This is Canela. She attempted some fancy porch...  
 24 You may not have known you needed to see this ...  
 25 This... is a Jubilant Antarctic House Bear. We...  
 26 This is Maya. She's very shy. Rarely leaves he...  
 27 This is Mingus. He's a wonderful father to his...  
 28 This is Derek. He's late for a dog meeting. 13...  
 29 This is Roscoe. Another pupper fallen victim t...  
 31 This is Waffles. His doggles are pupside down...  
 ...  
 2326 This is quite the dog. Gets really excited whe...  
 2327 This is a southern Vesuvius bumblegruff. Can d...



2328 Oh goodness. A super rare northeast Qdoba kang...  
 2329 Those are sunglasses and a jean jacket. 11/10 ...  
 2330 Unique dog here. Very small. Lives in containe...  
 2331 Here we have a mixed Asiago from the Galápagos...  
 2332 Look at this jokester thinking seat belt laws ...  
 2333 This is an extremely rare horned Parthenon. No...  
 2334 This is a funny dog. Weird toes. Won't come do...  
 2335 This is an Albanian 3 1/2 legged Episcopalian...  
 2336 Can take selfies 11/10 <https://t.co/ws2AMaNwPW>  
 2337 Very concerned about fellow dog trapped in com...  
 2338 Not familiar with this breed. No tail (weird)...  
 2339 Oh my. Here you are seeing an Adobe Setter giv...  
 2340 Can stand on stump for what seems like a while...  
 2341 This appears to be a Mongolian Presbyterian mi...  
 2342 Here we have a well-established sunblockerspan...  
 2343 Let's hope this flight isn't Malaysian (lol). ...  
 2344 Here we have a northern speckled Rhododendron...  
 2345 This is the happiest dog you will ever see. Ve...  
 2346 Here is the Rand Paul of retrievers folks! He'...  
 2347 My oh my. This is a rare blond Canadian terrie...  
 2348 Here is a Siberian heavily armored polar bear ...  
 2349 This is an odd dog. Hard on the outside but lo...  
 2350 This is a truly beautiful English Wilson Staff...  
 2351 Here we have a 1949 1st generation vulpix. Enj...  
 2352 This is a purebred Piers Morgan. Loves to Netf...  
 2353 Here is a very happy pup. Big fan of well-main...  
 2354 This is a western brown Mitsubishi terrier. Up...  
 2355 Here we have a Japanese Irish Setter. Lost eye...

	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	https://twitter.com/dog_rates/status/888554962...	13
20	https://twitter.com/dog_rates/status/888078434...	12
21	https://twitter.com/dog_rates/status/887705289...	13
22	https://twitter.com/dog_rates/status/887517139...	14
23	https://twitter.com/dog_rates/status/887473957...	13
24	https://twitter.com/dog_rates/status/887343217...	13
25	https://twitter.com/dog_rates/status/887101392...	12
26	https://twitter.com/dog_rates/status/886983233...	13
27	https://www.gofundme.com/mingusneedsus,https://...	13
28	https://twitter.com/dog_rates/status/886680336...	13
29	https://twitter.com/dog_rates/status/886366144...	12
31	https://twitter.com/dog_rates/status/886258384...	13
...	...	...
2326	https://twitter.com/dog_rates/status/666411507...	2
2327	https://twitter.com/dog_rates/status/666407126...	7
2328	https://twitter.com/dog_rates/status/666396247...	9
2329	https://twitter.com/dog_rates/status/666373753...	11
2330	https://twitter.com/dog_rates/status/666362758...	6
2331	https://twitter.com/dog_rates/status/666353288...	8
2332	https://twitter.com/dog_rates/status/666345417...	10
2333	https://twitter.com/dog_rates/status/666337882...	9
2334	https://twitter.com/dog_rates/status/666293911...	3
2335	https://twitter.com/dog_rates/status/666287406...	1
2336	https://twitter.com/dog_rates/status/666273097...	11
2337	https://twitter.com/dog_rates/status/666268910...	10
2338	https://twitter.com/dog_rates/status/666104133...	1
2339	https://twitter.com/dog_rates/status/666102155...	11
2340	https://twitter.com/dog_rates/status/666099513...	8
2341	https://twitter.com/dog_rates/status/666094000...	9
2342	https://twitter.com/dog_rates/status/666082916...	6
2343	https://twitter.com/dog_rates/status/666073100...	10
2344	https://twitter.com/dog_rates/status/666071193...	9
2345	https://twitter.com/dog_rates/status/666063827...	10
2346	https://twitter.com/dog_rates/status/666058600...	8
2347	https://twitter.com/dog_rates/status/666057090...	9
2348	https://twitter.com/dog_rates/status/666055525...	10
2349	https://twitter.com/dog_rates/status/666051853...	2
2350	https://twitter.com/dog_rates/status/666050758...	10
2351	https://twitter.com/dog_rates/status/666049248...	5
2352	https://twitter.com/dog_rates/status/666044226...	6
2353	https://twitter.com/dog_rates/status/666033412...	9
2354	https://twitter.com/dog_rates/status/666029285...	7
2355	https://twitter.com/dog_rates/status/666020888...	8

	rating_denominator	name	doggo	floofer	pupper	puppo
0	10	Phineas	None	None	None	None
1	10	Tilly	None	None	None	None
2	10	Archie	None	None	None	None

3	10	Darla	None	None	None	None
4	10	Franklin	None	None	None	None
5	10	None	None	None	None	None
6	10	Jax	None	None	None	None
7	10	None	None	None	None	None
8	10	Zoey	None	None	None	None
9	10	Cassie	doggo	None	None	None
10	10	Koda	None	None	None	None
11	10	Bruno	None	None	None	None
12	10	None	None	None	None	puppo
13	10	Ted	None	None	None	None
14	10	Stuart	None	None	None	puppo
15	10	Oliver	None	None	None	None
16	10	Jim	None	None	None	None
17	10	Zeke	None	None	None	None
18	10	Ralphus	None	None	None	None
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
31	10	Waffles	None	None	None	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

[2097 rows x 12 columns]

## Test

```
In [30]: df1_clean.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 12 columns):
tweet_id          2097 non-null int64
timestamp         2097 non-null datetime64[ns]
source            2097 non-null object
text              2097 non-null object
expanded_urls     2094 non-null object
rating_numerator  2097 non-null int64
rating_denominator 2097 non-null int64
name              2097 non-null object
doggo             2097 non-null object
floofer          2097 non-null object
pupper           2097 non-null object
puppo            2097 non-null object
dtypes: datetime64[ns](1), int64(3), object(8)
memory usage: 213.0+ KB
```

## Define

- 'df\_1' dataset : extract source of tweet from source column and categorize it.

## Code

```
In [31]: ## extract unique sources .
```

```
df1_clean['source'] = df1_clean['source'].map(lambda x: x.lower())
df1_clean['source'].unique()
```

```
Out[31]: array(['<a href="http://twitter.com/download/iphone" rel="nofollow">twitter for iphone',
                '<a href="http://twitter.com" rel="nofollow">twitter web client</a>'],
```

```
'<a href="http://vine.co" rel="nofollow">vine - make a scene</a>',
'<a href="https://about.twitter.com/products/tweetdeck" rel="nofollow">tweetdeck
```

```
In [32]: ## extract unique values string between '>' , '<' from column as source.
```

```
df1_clean['source'] = df1_clean['source'].str.split('>').str[1]
df1_clean['source'] = df1_clean['source'].str.split('<').str[0]
df1_clean['source'] = df1_clean['source'].astype('category')

df1_clean['source'].value_counts()
```

```
Out[32]: twitter for iphone      1964
vine - make a scene             91
twitter web client              31
tweetdeck                      11
Name: source, dtype: int64
```

## Test

```
In [33]: df1_clean['source'].value_counts()
```

```
Out[33]: twitter for iphone      1964
vine - make a scene             91
twitter web client              31
tweetdeck                      11
Name: source, dtype: int64
```

```
In [34]: df1_clean.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 12 columns):
tweet_id      2097 non-null int64
timestamp     2097 non-null datetime64[ns]
source        2097 non-null category
text          2097 non-null object
expanded_urls 2094 non-null object
rating_numerator 2097 non-null int64
rating_denominator 2097 non-null int64
name          2097 non-null object
doggo         2097 non-null object
floofer       2097 non-null object
pupper        2097 non-null object
puppo         2097 non-null object
dtypes: category(1), datetime64[ns](1), int64(3), object(7)
memory usage: 198.8+ KB
```

## Define

- 'df\_1' dataset : text column contain urls.

## Code

```
In [35]: ## replace url with space.
df1_clean['text'] = df1_clean['text'].str.replace(r'(https?:\\\/(?:www\.)?)[-a-zA-Z0-9@:
                                                ""')
```

## Test

```
In [36]: with pd.option_context('display.max_colwidth', None):
display(df1_clean.iloc[4])
```

```
-----

ValueError                                Traceback (most recent call last)

<ipython-input-36-9aa94b1e767a> in <module>()
----> 1 with pd.option_context('display.max_colwidth', None):
      2     display(df1_clean.iloc[4])

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in __enter__(self)
404
405     for pat, val in self.ops:
--> 406         _set_option(pat, val, silent=True)
407
408     def __exit__(self, *args):

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in _set_option(*args, **kwargs)
121     o = _get_registered_option(key)
122     if o and o.validator:
--> 123         o.validator(v)
124
125     # walk the nested dict

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in inner(x)
762     if type(x) != _type:
763         msg = "Value must have type '{typ!s}'"
--> 764         raise ValueError(msg.format(typ=_type))
765
766     return inner

ValueError: Value must have type '<class 'int'>'
```

```
In [37]: df1_clean.info()
```

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2097 entries, 0 to 2355
Data columns (total 12 columns):
tweet_id          2097 non-null int64
timestamp         2097 non-null datetime64[ns]
source            2097 non-null category
text              2097 non-null object
expanded_urls     2094 non-null object
rating_numerator  2097 non-null int64
rating_denominator 2097 non-null int64
name              2097 non-null object
doggo             2097 non-null object
floofer          2097 non-null object
pupper           2097 non-null object
puppo            2097 non-null object
dtypes: category(1), datetime64[ns](1), int64(3), object(7)
memory usage: 198.8+ KB

```

## Define

- 'df\_1' dataset : expanded\_urls have NaN values.

## Code

In [38]: *## not\_null values are the required rows so I used not\_null method to drop null rows.*

```
df1_clean= df1_clean[df1_clean['expanded_urls'].notnull()]
```

```
df1_clean
```

```

Out[38]:
   tweet_id          timestamp      source \
0  892420643555336193  2017-08-01 16:23:56  twitter for iphone
1  892177421306343426  2017-08-01 00:17:27  twitter for iphone
2  891815181378084864  2017-07-31 00:18:03  twitter for iphone
3  891689557279858688  2017-07-30 15:58:51  twitter for iphone
4  891327558926688256  2017-07-29 16:00:24  twitter for iphone
5  891087950875897856  2017-07-29 00:08:17  twitter for iphone
6  890971913173991426  2017-07-28 16:27:12  twitter for iphone
7  890729181411237888  2017-07-28 00:22:40  twitter for iphone
8  890609185150312448  2017-07-27 16:25:51  twitter for iphone
9  890240255349198849  2017-07-26 15:59:51  twitter for iphone
10 890006608113172480  2017-07-26 00:31:25  twitter for iphone
11 889880896479866881  2017-07-25 16:11:53  twitter for iphone
12 889665388333682689  2017-07-25 01:55:32  twitter for iphone
13 889638837579907072  2017-07-25 00:10:02  twitter for iphone
14 889531135344209921  2017-07-24 17:02:04  twitter for iphone
15 889278841981685760  2017-07-24 00:19:32  twitter for iphone

```

16	888917238123831296	2017-07-23	00:22:39	twitter for iphone
17	888804989199671297	2017-07-22	16:56:37	twitter for iphone
18	888554962724278272	2017-07-22	00:23:06	twitter for iphone
20	888078434458587136	2017-07-20	16:49:33	twitter for iphone
21	887705289381826560	2017-07-19	16:06:48	twitter for iphone
22	887517139158093824	2017-07-19	03:39:09	twitter for iphone
23	887473957103951883	2017-07-19	00:47:34	twitter for iphone
24	887343217045368832	2017-07-18	16:08:03	twitter for iphone
25	887101392804085760	2017-07-18	00:07:08	twitter for iphone
26	886983233522544640	2017-07-17	16:17:36	twitter for iphone
27	886736880519319552	2017-07-16	23:58:41	twitter for iphone
28	886680336477933568	2017-07-16	20:14:00	twitter for iphone
29	886366144734445568	2017-07-15	23:25:31	twitter for iphone
31	886258384151887873	2017-07-15	16:17:19	twitter for iphone
...	...	...	...	...
2326	666411507551481857	2015-11-17	00:24:19	twitter for iphone
2327	666407126856765440	2015-11-17	00:06:54	twitter for iphone
2328	666396247373291520	2015-11-16	23:23:41	twitter for iphone
2329	666373753744588802	2015-11-16	21:54:18	twitter for iphone
2330	666362758909284353	2015-11-16	21:10:36	twitter for iphone
2331	666353288456101888	2015-11-16	20:32:58	twitter for iphone
2332	666345417576210432	2015-11-16	20:01:42	twitter for iphone
2333	666337882303524864	2015-11-16	19:31:45	twitter for iphone
2334	666293911632134144	2015-11-16	16:37:02	twitter for iphone
2335	666287406224695296	2015-11-16	16:11:11	twitter for iphone
2336	666273097616637952	2015-11-16	15:14:19	twitter for iphone
2337	666268910803644416	2015-11-16	14:57:41	twitter for iphone
2338	666104133288665088	2015-11-16	04:02:55	twitter for iphone
2339	666102155909144576	2015-11-16	03:55:04	twitter for iphone
2340	666099513787052032	2015-11-16	03:44:34	twitter for iphone
2341	666094000022159362	2015-11-16	03:22:39	twitter for iphone
2342	666082916733198337	2015-11-16	02:38:37	twitter for iphone
2343	666073100786774016	2015-11-16	01:59:36	twitter for iphone
2344	666071193221509120	2015-11-16	01:52:02	twitter for iphone
2345	666063827256086533	2015-11-16	01:22:45	twitter for iphone
2346	666058600524156928	2015-11-16	01:01:59	twitter for iphone
2347	666057090499244032	2015-11-16	00:55:59	twitter for iphone
2348	666055525042405380	2015-11-16	00:49:46	twitter for iphone
2349	666051853826850816	2015-11-16	00:35:11	twitter for iphone
2350	666050758794694657	2015-11-16	00:30:50	twitter for iphone
2351	666049248165822465	2015-11-16	00:24:50	twitter for iphone
2352	666044226329800704	2015-11-16	00:04:52	twitter for iphone
2353	666033412701032449	2015-11-15	23:21:54	twitter for iphone
2354	666029285002620928	2015-11-15	23:05:30	twitter for iphone
2355	666020888022790149	2015-11-15	22:32:08	twitter for iphone

text \

0 This is Phineas. He's a mystical boy. Only eve...



1 This is Tilly. She's just checking pup on you...  
 2 This is Archie. He is a rare Norwegian Pouncin...  
 3 This is Darla. She commenced a snooze mid meal...  
 4 This is Franklin. He would like you to stop ca...  
 5 Here we have a majestic great white breaching ...  
 6 Meet Jax. He enjoys ice cream so much he gets ...  
 7 When you watch your owner call another dog a g...  
 8 This is Zoey. She doesn't want to be one of th...  
 9 This is Cassie. She is a college pup. Studying...  
 10 This is Koda. He is a South Australian decksha...  
 11 This is Bruno. He is a service shark. Only get...  
 12 Here's a puppo that seems to be on the fence a...  
 13 This is Ted. He does his best. Sometimes that'...  
 14 This is Stuart. He's sporting his favorite fan...  
 15 This is Oliver. You're witnessing one of his m...  
 16 This is Jim. He found a fren. Taught him how t...  
 17 This is Zeke. He has a new stick. Very proud o...  
 18 This is Ralphus. He's powering up. Attempting ...  
 20 This is Gerald. He was just told he didn't get...  
 21 This is Jeffrey. He has a monopoly on the pool...  
 22 I've yet to rate a Venezuelan Hover Wiener. Th...  
 23 This is Canela. She attempted some fancy porch...  
 24 You may not have known you needed to see this ...  
 25 This... is a Jubilant Antarctic House Bear. We...  
 26 This is Maya. She's very shy. Rarely leaves he...  
 27 This is Mingus. He's a wonderful father to his...  
 28 This is Derek. He's late for a dog meeting. 13...  
 29 This is Roscoe. Another pupper fallen victim t...  
 31 This is Waffles. His doggles are pupside down...  
 ...  
 2326 This is quite the dog. Gets really excited whe...  
 2327 This is a southern Vesuvius bumblegruff. Can d...  
 2328 Oh goodness. A super rare northeast Qdoba kang...  
 2329 Those are sunglasses and a jean jacket. 11/10 ...  
 2330 Unique dog here. Very small. Lives in containe...  
 2331 Here we have a mixed Asiago from the Galápagos...  
 2332 Look at this jokester thinking seat belt laws ...  
 2333 This is an extremely rare horned Parthenon. No...  
 2334 This is a funny dog. Weird toes. Won't come do...  
 2335 This is an Albanian 3 1/2 legged Episcopalian...  
 2336 Can take selfies 11/10  
 2337 Very concerned about fellow dog trapped in com...  
 2338 Not familiar with this breed. No tail (weird)...  
 2339 Oh my. Here you are seeing an Adobe Setter giv...  
 2340 Can stand on stump for what seems like a while...  
 2341 This appears to be a Mongolian Presbyterian mi...  
 2342 Here we have a well-established sunblockerspan...  
 2343 Let's hope this flight isn't Malaysian (lol). ...

2344 Here we have a northern speckled Rhododendron...  
 2345 This is the happiest dog you will ever see. Ve...  
 2346 Here is the Rand Paul of retrievers folks! He'...  
 2347 My oh my. This is a rare blond Canadian terrie...  
 2348 Here is a Siberian heavily armored polar bear ...  
 2349 This is an odd dog. Hard on the outside but lo...  
 2350 This is a truly beautiful English Wilson Staff...  
 2351 Here we have a 1949 1st generation vulpix. Enj...  
 2352 This is a purebred Piers Morgan. Loves to Netf...  
 2353 Here is a very happy pup. Big fan of well-main...  
 2354 This is a western brown Mitsubishi terrier. Up...  
 2355 Here we have a Japanese Irish Setter. Lost eye...

	expanded_urls	rating_numerator \
0	https://twitter.com/dog_rates/status/892420643...	13
1	https://twitter.com/dog_rates/status/892177421...	13
2	https://twitter.com/dog_rates/status/891815181...	12
3	https://twitter.com/dog_rates/status/891689557...	13
4	https://twitter.com/dog_rates/status/891327558...	12
5	https://twitter.com/dog_rates/status/891087950...	13
6	https://gofundme.com/ydvmve-surgery-for-jax,ht...	13
7	https://twitter.com/dog_rates/status/890729181...	13
8	https://twitter.com/dog_rates/status/890609185...	13
9	https://twitter.com/dog_rates/status/890240255...	14
10	https://twitter.com/dog_rates/status/890006608...	13
11	https://twitter.com/dog_rates/status/889880896...	13
12	https://twitter.com/dog_rates/status/889665388...	13
13	https://twitter.com/dog_rates/status/889638837...	12
14	https://twitter.com/dog_rates/status/889531135...	13
15	https://twitter.com/dog_rates/status/889278841...	13
16	https://twitter.com/dog_rates/status/888917238...	12
17	https://twitter.com/dog_rates/status/888804989...	13
18	https://twitter.com/dog_rates/status/888554962...	13
20	https://twitter.com/dog_rates/status/888078434...	12
21	https://twitter.com/dog_rates/status/887705289...	13
22	https://twitter.com/dog_rates/status/887517139...	14
23	https://twitter.com/dog_rates/status/887473957...	13
24	https://twitter.com/dog_rates/status/887343217...	13
25	https://twitter.com/dog_rates/status/887101392...	12
26	https://twitter.com/dog_rates/status/886983233...	13
27	https://www.gofundme.com/mingusneedsus,https:/...	13
28	https://twitter.com/dog_rates/status/886680336...	13
29	https://twitter.com/dog_rates/status/886366144...	12
31	https://twitter.com/dog_rates/status/886258384...	13
...	...	...
2326	https://twitter.com/dog_rates/status/666411507...	2
2327	https://twitter.com/dog_rates/status/666407126...	7
2328	https://twitter.com/dog_rates/status/666396247...	9

2329	<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

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
31	10	Waffles	None	None	None	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

[2094 rows x 12 columns]

## Test

In [39]: df1\_clean.info()

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2094 entries, 0 to 2355
Data columns (total 12 columns):
tweet_id          2094 non-null int64
timestamp         2094 non-null datetime64[ns]
source            2094 non-null category
text              2094 non-null object
expanded_urls     2094 non-null object
rating_numerator  2094 non-null int64
rating_denominator 2094 non-null int64
name              2094 non-null object
doggo             2094 non-null object
floofer          2094 non-null object
pupper           2094 non-null object
puppo            2094 non-null object
dtypes: category(1), datetime64[ns](1), int64(3), object(7)
memory usage: 198.5+ KB

```

## Define

- 'df\_1' dataset : some expanded\_urls contain duplicated urls in the same value.

## Code

```

In [40]: ## divid dataset to 2 datasets by expanded_urls depend on contain ',' or not .
         c1 = df1_clean.copy()
         nc2 = df1_clean.copy()

         c1 = c1[c1['expanded_urls'].str.contains(',')]
         nc2 = nc2[~nc2['expanded_urls'].str.contains(',')]

         ## extract one url from expanded_urls column in dataset contain ',' .
         c1['expanded_urls'] = c1['expanded_urls'].str.split(',', expand=True)[1]

         ## concat 2 new datasets.
         df1_clean = pd.concat([c1, nc2])

```

## Test

```

In [41]: with pd.option_context('display.max_colwidth', None):
         display(df1_clean.sample(50))

```

```

-----
ValueError                                Traceback (most recent call last)

```

```

<ipython-input-41-26a720239355> in <module>()
----> 1 with pd.option_context('display.max_colwidth', None):
      2     display(df1_clean.sample(50))

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in __enter__(self)
404
405         for pat, val in self.ops:
--> 406             _set_option(pat, val, silent=True)
407
408     def __exit__(self, *args):

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in _set_option(*args, **kwargs)
121         o = _get_registered_option(key)
122         if o and o.validator:
--> 123             o.validator(v)
124
125         # walk the nested dict

/opt/conda/lib/python3.6/site-packages/pandas/core/config.py in inner(x)
762         if type(x) != _type:
763             msg = "Value must have type '{typ!s}'"
--> 764             raise ValueError(msg.format(typ=_type))
765
766     return inner

ValueError: Value must have type '<class 'int'>'

```

In [42]: df1\_clean.info()

```

<class 'pandas.core.frame.DataFrame'>
Int64Index: 2094 entries, 4 to 2355
Data columns (total 12 columns):
tweet_id          2094 non-null int64
timestamp         2094 non-null datetime64[ns]
source            2094 non-null category
text              2094 non-null object
expanded_urls     2094 non-null object
rating_numerator  2094 non-null int64
rating_denominator 2094 non-null int64
name              2094 non-null object
doggo             2094 non-null object
floofer           2094 non-null object
pupper           2094 non-null object

```

```
puppo                2094 non-null object
dtypes: category(1), datetime64[ns](1), int64(3), object(7)
memory usage: 198.5+ KB
```

## Define

- 'df\_1' dataset : some rating\_denominator value == accuracy issue.

## Code

```
In [43]: ## rating_denominator value should be 10 acc. to unique classification , so I will drop
```

```
df1_clean = df1_clean.loc[df1_clean['rating_denominator'] == 10]
```

## Test

```
In [44]: df1_clean.describe()
```

```
Out[44]:
```

	tweet_id	rating_numerator	rating_denominator
count	2.077000e+03	2077.000000	2077.0
mean	7.366544e+17	11.699567	10.0
std	6.725115e+16	39.845599	0.0
min	6.660209e+17	0.000000	10.0
25%	6.766175e+17	10.000000	10.0
50%	7.095662e+17	11.000000	10.0
75%	7.880396e+17	12.000000	10.0
max	8.924206e+17	1776.000000	10.0

```
In [45]: df1_clean.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 2077 entries, 4 to 2355
Data columns (total 12 columns):
tweet_id                2077 non-null int64
timestamp               2077 non-null datetime64[ns]
source                  2077 non-null category
text                    2077 non-null object
expanded_urls           2077 non-null object
rating_numerator        2077 non-null int64
rating_denominator      2077 non-null int64
name                    2077 non-null object
doggo                   2077 non-null object
floofer                 2077 non-null object
pupper                 2077 non-null object
puppo                   2077 non-null object
dtypes: category(1), datetime64[ns](1), int64(3), object(7)
memory usage: 196.9+ KB
```

## Define

- 'df\_1' dataset : some rating\_numerator value not == validity issue.

## Code

```
In [46]: ## rating_numerators less than or equal 5 most of them are not dog ratings , I will drop
df1_clean = df1_clean [df1_clean['rating_numerator'] >= 6]

## rating_numerators more than 100 has consistency issue , I will drop them .
df1_clean = df1_clean [df1_clean['rating_numerator'] <= 100]

## some rating_numerators between 5 and 100 are flots .
## I will extract the rating_numerator again from text and convert it to float.
df1_clean['rating_numerator'] = df1_clean['text'].str.extract('(\d+\.?\d?\d?)\\/\d{1,3}')
```

## Test

```
In [47]: df1_clean.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 1993 entries, 4 to 2355
Data columns (total 12 columns):
tweet_id          1993 non-null int64
timestamp         1993 non-null datetime64[ns]
source            1993 non-null category
text              1993 non-null object
expanded_urls     1993 non-null object
rating_numerator  1993 non-null float64
rating_denominator 1993 non-null int64
name              1993 non-null object
doggo             1993 non-null object
floofer          1993 non-null object
pupper           1993 non-null object
puppo            1993 non-null object
dtypes: category(1), datetime64[ns](1), float64(1), int64(2), object(7)
memory usage: 189.0+ KB
```

```
In [48]: df1_clean.describe()
```

```
Out[48]:
```

	tweet_id	rating_numerator	rating_denominator
count	1.993000e+03	1993.000000	1993.0
mean	7.386751e+17	10.887747	10.0
std	6.747765e+16	1.660368	0.0
min	6.660209e+17	6.000000	10.0
25%	6.776740e+17	10.000000	10.0
50%	7.127178e+17	11.000000	10.0
75%	7.909874e+17	12.000000	10.0
max	8.924206e+17	14.000000	10.0



## Define

- 'df\_1' dataset : doggo,floofer,pupper,puppo columns have 'none' for missing values.

## Code

```
In [49]: ## replace none by "" , combine columns doggo,floofer,pupper,puppo as fancy_class.
df1_clean.iloc[:, -4: ] = df1_clean.iloc[:, -4: ].replace('None','')
df1_clean['fancy_class'] = df1_clean.doggo + df1_clean.floofer + df1_clean.pupper + df1

## drop columns doggo,floofer,pupper,puppo , replace "" by NaN .
df1_clean = df1_clean.drop(columns=['doggo','floofer','pupper','puppo'] , axis = 1)
df1_clean.iloc[:, -1] = df1_clean.iloc[:, -1].replace('',np.nan)
df1_clean
```

```
Out[49]:
```

	tweet_id	timestamp	source \
4	891327558926688256	2017-07-29 16:00:24	twitter for iphone
6	890971913173991426	2017-07-28 16:27:12	twitter for iphone
7	890729181411237888	2017-07-28 00:22:40	twitter for iphone
10	890006608113172480	2017-07-26 00:31:25	twitter for iphone
13	889638837579907072	2017-07-25 00:10:02	twitter for iphone
17	888804989199671297	2017-07-22 16:56:37	twitter for iphone
18	888554962724278272	2017-07-22 00:23:06	twitter for iphone
20	888078434458587136	2017-07-20 16:49:33	twitter for iphone
23	887473957103951883	2017-07-19 00:47:34	twitter for iphone
26	886983233522544640	2017-07-17 16:17:36	twitter for iphone
27	886736880519319552	2017-07-16 23:58:41	twitter for iphone
29	886366144734445568	2017-07-15 23:25:31	twitter for iphone
37	885167619883638784	2017-07-12 16:03:00	twitter for iphone
39	884876753390489601	2017-07-11 20:47:12	twitter for iphone
47	883117836046086144	2017-07-07 00:17:54	twitter for iphone
48	882992080364220416	2017-07-06 15:58:11	twitter for iphone
51	882268110199369728	2017-07-04 16:01:23	twitter for iphone
52	882045870035918850	2017-07-04 01:18:17	twitter for iphone
60	880465832366813184	2017-06-29 16:39:47	twitter for iphone
61	880221127280381952	2017-06-29 00:27:25	twitter for iphone
63	879862464715927552	2017-06-28 00:42:13	twitter for iphone
71	878776093423087618	2017-06-25 00:45:22	twitter for iphone
75	878281511006478336	2017-06-23 16:00:04	twitter for iphone
76	878057613040115712	2017-06-23 01:10:23	twitter for iphone
77	877736472329191424	2017-06-22 03:54:17	twitter for iphone
80	877316821321428993	2017-06-21 00:06:44	twitter for iphone
81	877201837425926144	2017-06-20 16:29:50	twitter for iphone
82	876838120628539392	2017-06-19 16:24:33	twitter for iphone
89	875021211251597312	2017-06-14 16:04:48	twitter for iphone
94	874012996292530176	2017-06-11 21:18:31	twitter for iphone
...	...	...	...
2320	666437273139982337	2015-11-17 02:06:42	twitter for iphone
2321	666435652385423360	2015-11-17 02:00:15	twitter for iphone

2322	666430724426358785	2015-11-17	01:40:41	twitter for iphone
2323	666428276349472768	2015-11-17	01:30:57	twitter for iphone
2324	666421158376562688	2015-11-17	01:02:40	twitter for iphone
2325	666418789513326592	2015-11-17	00:53:15	twitter for iphone
2327	666407126856765440	2015-11-17	00:06:54	twitter for iphone
2328	666396247373291520	2015-11-16	23:23:41	twitter for iphone
2329	666373753744588802	2015-11-16	21:54:18	twitter for iphone
2330	666362758909284353	2015-11-16	21:10:36	twitter for iphone
2331	666353288456101888	2015-11-16	20:32:58	twitter for iphone
2332	666345417576210432	2015-11-16	20:01:42	twitter for iphone
2333	666337882303524864	2015-11-16	19:31:45	twitter for iphone
2336	666273097616637952	2015-11-16	15:14:19	twitter for iphone
2337	666268910803644416	2015-11-16	14:57:41	twitter for iphone
2339	666102155909144576	2015-11-16	03:55:04	twitter for iphone
2340	666099513787052032	2015-11-16	03:44:34	twitter for iphone
2341	666094000022159362	2015-11-16	03:22:39	twitter for iphone
2342	666082916733198337	2015-11-16	02:38:37	twitter for iphone
2343	666073100786774016	2015-11-16	01:59:36	twitter for iphone
2344	666071193221509120	2015-11-16	01:52:02	twitter for iphone
2345	666063827256086533	2015-11-16	01:22:45	twitter for iphone
2346	666058600524156928	2015-11-16	01:01:59	twitter for iphone
2347	666057090499244032	2015-11-16	00:55:59	twitter for iphone
2348	666055525042405380	2015-11-16	00:49:46	twitter for iphone
2350	666050758794694657	2015-11-16	00:30:50	twitter for iphone
2352	666044226329800704	2015-11-16	00:04:52	twitter for iphone
2353	666033412701032449	2015-11-15	23:21:54	twitter for iphone
2354	666029285002620928	2015-11-15	23:05:30	twitter for iphone
2355	666020888022790149	2015-11-15	22:32:08	twitter for iphone

text \

4	This is Franklin. He would like you to stop ca...
6	Meet Jax. He enjoys ice cream so much he gets ...
7	When you watch your owner call another dog a g...
10	This is Koda. He is a South Australian decksha...
13	This is Ted. He does his best. Sometimes that'...
17	This is Zeke. He has a new stick. Very proud o...
18	This is Ralphus. He's powering up. Attempting ...
20	This is Gerald. He was just told he didn't get...
23	This is Canela. She attempted some fancy porch...
26	This is Maya. She's very shy. Rarely leaves he...
27	This is Mingus. He's a wonderful father to his...
29	This is Roscoe. Another pupper fallen victim t...
37	Here we have a corgi undercover as a malamute...
39	This is Lola. It's her first time outside. Mus...
47	Please only send dogs. We don't rate mechanics...
48	This is Rusty. He wasn't ready for the first p...
51	This is Alf. You're witnessing his first wate...
52	This is Koko. Her owner, inspired by Barney, r...

60 This is Bella. She had her first beach experie...  
 61 Meet Jesse. He's a Fetty Woof. His tongue ejec...  
 63 This is Romeo. He would like to do an entrance...  
 71 This is Snoopy. He's a proud #PrideMonthPuppo...  
 75 Meet Shadow. In an attempt to reach maximum zo...  
 76 This is Emmy. She was adopted today. Massive r...  
 77 This is Aja. She was just told she's a good do...  
 80 Meet Dante. At first he wasn't a fan of his ne...  
 81 This is Nelly. He graduated with his dogtorate...  
 82 This is Ginger. She's having a ruff Monday. To...  
 89 Guys please stop sending pictures without any ...  
 94 This is Sebastian. He can't see all the colors...  
 ...  
 2320 Here we see a lone northeastern Cumberbatch. H...  
 2321 "Can you behave? You're ruining my wedding day...  
 2322 Oh boy what a pup! Sunglasses take this one to...  
 2323 Here we have an Austrian Pulitzer. Collectors ...  
 2324 \*internally screaming\* 12/10  
 2325 This is Walter. He is an Alaskan Terrapin. Lov...  
 2327 This is a southern Vesuvius bumblegruff. Can d...  
 2328 Oh goodness. A super rare northeast Qdoba kang...  
 2329 Those are sunglasses and a jean jacket. 11/10 ...  
 2330 Unique dog here. Very small. Lives in containe...  
 2331 Here we have a mixed Asiago from the Galápagos...  
 2332 Look at this jokester thinking seat belt laws ...  
 2333 This is an extremely rare horned Parthenon. No...  
 2336 Can take selfies 11/10  
 2337 Very concerned about fellow dog trapped in com...  
 2339 Oh my. Here you are seeing an Adobe Setter giv...  
 2340 Can stand on stump for what seems like a while...  
 2341 This appears to be a Mongolian Presbyterian mi...  
 2342 Here we have a well-established sunblockerspan...  
 2343 Let's hope this flight isn't Malaysian (lol). ...  
 2344 Here we have a northern speckled Rhododendron...  
 2345 This is the happiest dog you will ever see. Ve...  
 2346 Here is the Rand Paul of retrievers folks! He'...  
 2347 My oh my. This is a rare blond Canadian terrie...  
 2348 Here is a Siberian heavily armored polar bear ...  
 2350 This is a truly beautiful English Wilson Staff...  
 2352 This is a purebred Piers Morgan. Loves to Netf...  
 2353 Here is a very happy pup. Big fan of well-main...  
 2354 This is a western brown Mitsubishi terrier. Up...  
 2355 Here we have a Japanese Irish Setter. Lost eye...

	expanded_urls	rating_numerator \
4	<a href="https://twitter.com/dog_rates/status/891327558...">https://twitter.com/dog_rates/status/891327558...</a>	12.0
6	<a href="https://twitter.com/dog_rates/status/890971913...">https://twitter.com/dog_rates/status/890971913...</a>	13.0
7	<a href="https://twitter.com/dog_rates/status/890729181...">https://twitter.com/dog_rates/status/890729181...</a>	13.0

10	<a href="https://twitter.com/dog_rates/status/890006608...">https://twitter.com/dog_rates/status/890006608...</a>	13.0
13	<a href="https://twitter.com/dog_rates/status/889638837...">https://twitter.com/dog_rates/status/889638837...</a>	12.0
17	<a href="https://twitter.com/dog_rates/status/888804989...">https://twitter.com/dog_rates/status/888804989...</a>	13.0
18	<a href="https://twitter.com/dog_rates/status/888554962...">https://twitter.com/dog_rates/status/888554962...</a>	13.0
20	<a href="https://twitter.com/dog_rates/status/888078434...">https://twitter.com/dog_rates/status/888078434...</a>	12.0
23	<a href="https://twitter.com/dog_rates/status/887473957...">https://twitter.com/dog_rates/status/887473957...</a>	13.0
26	<a href="https://twitter.com/dog_rates/status/886983233...">https://twitter.com/dog_rates/status/886983233...</a>	13.0
27	<a href="https://twitter.com/dog_rates/status/886736880...">https://twitter.com/dog_rates/status/886736880...</a>	13.0
29	<a href="https://twitter.com/dog_rates/status/886366144...">https://twitter.com/dog_rates/status/886366144...</a>	12.0
37	<a href="https://twitter.com/dog_rates/status/885167619...">https://twitter.com/dog_rates/status/885167619...</a>	13.0
39	<a href="https://twitter.com/dog_rates/status/884876753...">https://twitter.com/dog_rates/status/884876753...</a>	13.0
47	<a href="https://twitter.com/dog_rates/status/883117836...">https://twitter.com/dog_rates/status/883117836...</a>	13.0
48	<a href="https://twitter.com/dog_rates/status/882992080...">https://twitter.com/dog_rates/status/882992080...</a>	13.0
51	<a href="https://twitter.com/dog_rates/status/882268110...">https://twitter.com/dog_rates/status/882268110...</a>	13.0
52	<a href="https://twitter.com/dog_rates/status/882045870...">https://twitter.com/dog_rates/status/882045870...</a>	13.0
60	<a href="https://twitter.com/dog_rates/status/880465832...">https://twitter.com/dog_rates/status/880465832...</a>	12.0
61	<a href="https://twitter.com/dog_rates/status/880221127...">https://twitter.com/dog_rates/status/880221127...</a>	12.0
63	<a href="https://twitter.com/dog_rates/status/879862464...">https://twitter.com/dog_rates/status/879862464...</a>	13.0
71	<a href="https://twitter.com/dog_rates/status/878776093...">https://twitter.com/dog_rates/status/878776093...</a>	13.0
75	<a href="https://twitter.com/dog_rates/status/878281511...">https://twitter.com/dog_rates/status/878281511...</a>	13.0
76	<a href="https://twitter.com/dog_rates/status/878057613...">https://twitter.com/dog_rates/status/878057613...</a>	14.0
77	<a href="https://twitter.com/dog_rates/status/877736472...">https://twitter.com/dog_rates/status/877736472...</a>	13.0
80	<a href="https://twitter.com/dog_rates/status/877316821...">https://twitter.com/dog_rates/status/877316821...</a>	13.0
81	<a href="https://twitter.com/dog_rates/status/877201837...">https://twitter.com/dog_rates/status/877201837...</a>	12.0
82	<a href="https://twitter.com/dog_rates/status/876838120...">https://twitter.com/dog_rates/status/876838120...</a>	12.0
89	<a href="https://twitter.com/dog_rates/status/875021211...">https://twitter.com/dog_rates/status/875021211...</a>	12.0
94	<a href="https://twitter.com/dog_rates/status/874012996...">https://twitter.com/dog_rates/status/874012996...</a>	13.0
...	...	...
2320	<a href="https://twitter.com/dog_rates/status/666437273...">https://twitter.com/dog_rates/status/666437273...</a>	7.0
2321	<a href="https://twitter.com/dog_rates/status/666435652...">https://twitter.com/dog_rates/status/666435652...</a>	10.0
2322	<a href="https://twitter.com/dog_rates/status/666430724...">https://twitter.com/dog_rates/status/666430724...</a>	6.0
2323	<a href="https://twitter.com/dog_rates/status/666428276...">https://twitter.com/dog_rates/status/666428276...</a>	7.0
2324	<a href="https://twitter.com/dog_rates/status/666421158...">https://twitter.com/dog_rates/status/666421158...</a>	12.0
2325	<a href="https://twitter.com/dog_rates/status/666418789...">https://twitter.com/dog_rates/status/666418789...</a>	10.0
2327	<a href="https://twitter.com/dog_rates/status/666407126...">https://twitter.com/dog_rates/status/666407126...</a>	7.0
2328	<a href="https://twitter.com/dog_rates/status/666396247...">https://twitter.com/dog_rates/status/666396247...</a>	9.0
2329	<a href="https://twitter.com/dog_rates/status/666373753...">https://twitter.com/dog_rates/status/666373753...</a>	11.0
2330	<a href="https://twitter.com/dog_rates/status/666362758...">https://twitter.com/dog_rates/status/666362758...</a>	6.0
2331	<a href="https://twitter.com/dog_rates/status/666353288...">https://twitter.com/dog_rates/status/666353288...</a>	8.0
2332	<a href="https://twitter.com/dog_rates/status/666345417...">https://twitter.com/dog_rates/status/666345417...</a>	10.0
2333	<a href="https://twitter.com/dog_rates/status/666337882...">https://twitter.com/dog_rates/status/666337882...</a>	9.0
2336	<a href="https://twitter.com/dog_rates/status/666273097...">https://twitter.com/dog_rates/status/666273097...</a>	11.0
2337	<a href="https://twitter.com/dog_rates/status/666268910...">https://twitter.com/dog_rates/status/666268910...</a>	10.0
2339	<a href="https://twitter.com/dog_rates/status/666102155...">https://twitter.com/dog_rates/status/666102155...</a>	11.0
2340	<a href="https://twitter.com/dog_rates/status/666099513...">https://twitter.com/dog_rates/status/666099513...</a>	8.0
2341	<a href="https://twitter.com/dog_rates/status/666094000...">https://twitter.com/dog_rates/status/666094000...</a>	9.0
2342	<a href="https://twitter.com/dog_rates/status/666082916...">https://twitter.com/dog_rates/status/666082916...</a>	6.0
2343	<a href="https://twitter.com/dog_rates/status/666073100...">https://twitter.com/dog_rates/status/666073100...</a>	10.0

2344	<a href="https://twitter.com/dog_rates/status/666071193...">https://twitter.com/dog_rates/status/666071193...</a>	9.0
2345	<a href="https://twitter.com/dog_rates/status/666063827...">https://twitter.com/dog_rates/status/666063827...</a>	10.0
2346	<a href="https://twitter.com/dog_rates/status/666058600...">https://twitter.com/dog_rates/status/666058600...</a>	8.0
2347	<a href="https://twitter.com/dog_rates/status/666057090...">https://twitter.com/dog_rates/status/666057090...</a>	9.0
2348	<a href="https://twitter.com/dog_rates/status/666055525...">https://twitter.com/dog_rates/status/666055525...</a>	10.0
2350	<a href="https://twitter.com/dog_rates/status/666050758...">https://twitter.com/dog_rates/status/666050758...</a>	10.0
2352	<a href="https://twitter.com/dog_rates/status/666044226...">https://twitter.com/dog_rates/status/666044226...</a>	6.0
2353	<a href="https://twitter.com/dog_rates/status/666033412...">https://twitter.com/dog_rates/status/666033412...</a>	9.0
2354	<a href="https://twitter.com/dog_rates/status/666029285...">https://twitter.com/dog_rates/status/666029285...</a>	7.0
2355	<a href="https://twitter.com/dog_rates/status/666020888...">https://twitter.com/dog_rates/status/666020888...</a>	8.0

	rating_denominator	name	fancy_class
4	10	Franklin	NaN
6	10	Jax	NaN
7	10	None	NaN
10	10	Koda	NaN
13	10	Ted	NaN
17	10	Zeke	NaN
18	10	Ralphus	NaN
20	10	Gerald	NaN
23	10	Canela	NaN
26	10	Maya	NaN
27	10	Mingus	NaN
29	10	Roscoe	pupper
37	10	None	NaN
39	10	Lola	NaN
47	10	None	NaN
48	10	Rusty	NaN
51	10	Alfy	NaN
52	10	Koko	NaN
60	10	Bella	NaN
61	10	Jesse	NaN
63	10	Romeo	NaN
71	10	Snoopy	puppo
75	10	Shadow	NaN
76	10	Emmy	NaN
77	10	Aja	NaN
80	10	Dante	NaN
81	10	Nelly	NaN
82	10	Ginger	pupper
89	10	None	NaN
94	10	Sebastian	puppo
...	...	...	...
2320	10	None	NaN
2321	10	None	NaN
2322	10	None	NaN
2323	10	None	NaN
2324	10	None	NaN

2325	10	Walter	NaN
2327	10	a	NaN
2328	10	None	NaN
2329	10	None	NaN
2330	10	None	NaN
2331	10	None	NaN
2332	10	None	NaN
2333	10	an	NaN
2336	10	None	NaN
2337	10	None	NaN
2339	10	None	NaN
2340	10	None	NaN
2341	10	None	NaN
2342	10	None	NaN
2343	10	None	NaN
2344	10	None	NaN
2345	10	the	NaN
2346	10	the	NaN
2347	10	a	NaN
2348	10	a	NaN
2350	10	a	NaN
2352	10	a	NaN
2353	10	a	NaN
2354	10	a	NaN
2355	10	None	NaN

[1993 rows x 9 columns]

## Test

In [50]: df1\_clean.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1993 entries, 4 to 2355
Data columns (total 9 columns):
tweet_id          1993 non-null int64
timestamp         1993 non-null datetime64[ns]
source            1993 non-null category
text              1993 non-null object
expanded_urls     1993 non-null object
rating_numerator  1993 non-null float64
rating_denominator 1993 non-null int64
name              1993 non-null object
fancy_class       332 non-null object
dtypes: category(1), datetime64[ns](1), float64(1), int64(2), object(4)
memory usage: 142.3+ KB
```

## Define

- 'df\_1' dataset : some name column values == validity issue.

## Code

```
In [51]: ## replace 'a' , 'an' in name columns by pattern in text .
pattern_2 = re.compile(r'(?:(?:name(?:d)?)\s{1})(?:is\s)?([A-Za-z]+)')
for index, row in df1_clean.iterrows():
    try:
        if row['name'] == "a":
            c_name = re.findall(pattern_2, row['text'])[0]
            df1_clean.loc[index, 'name'] = df1_clean.loc[index, 'name'].replace('a', c_name)
        elif row['name'] == 'an':
            c_name = re.findall(pattern_2, row['text'])[0]
            df1_clean.loc[index, 'name'] = df1_clean.loc[index, 'name'].replace('an', c_name)
    except IndexError:
        df1_clean.loc[index, 'name'] = np.nan
```

## Test

```
In [52]: df1_clean.name.value_counts(dropna = False)
```

```
Out[52]: None      545
NaN             32
Charlie         11
Lucy            11
Cooper          10
Oliver          10
Tucker          9
Penny           9
Lola             8
Winston         8
the             8
Sadie           8
Daisy           7
Toby            7
Bailey          6
Stanley         6
Koda            6
Jax             6
Oscar           6
Bo              6
Louis           5
Bentley         5
Milo            5
Bella           5
Leo             5
Chester         5
```

Rusty	5
Scout	5
Buddy	5
Boomer	4
...	
Herb	1
Chuk	1
Comet	1
Brownie	1
Ed	1
Dallas	1
Karl	1
Tango	1
Ole	1
Lillie	1
Rizzo	1
Nida	1
Fwed	1
Paull	1
Timofy	1
Patrick	1
Tobi	1
Sobe	1
Mabel	1
Tiger	1
Theo	1
Clarq	1
Einstein	1
Rufio	1
Jersey	1
Malikai	1
Remy	1
Ziva	1
Callie	1
Vinnie	1

Name: name, Length: 955, dtype: int64

```
In [53]: df1_clean.name.unique()
```

```
Out[53]: array(['Franklin', 'Jax', 'None', 'Koda', 'Ted', 'Zeke', 'Ralphus',
                'Gerald', 'Canela', 'Maya', 'Mingus', 'Roscoe', 'Lola', 'Rusty',
                'Alfy', 'Koko', 'Bella', 'Jesse', 'Romeo', 'Snoopy', 'Shadow',
                'Emmy', 'Aja', 'Dante', 'Nelly', 'Ginger', 'Sebastian', 'Sierra',
                'Harry', 'Kody', 'Lassie', 'Rover', 'Napolean', 'Cody', 'Rumble',
                'Clifford', 'Scout', 'Cooper', 'Harold', 'Lili', 'Jamesy', 'Coco',
                'Sammy', 'Meatball', 'Paisley', 'Albus', 'Quinn', 'Zooey', 'Jersey',
                'Lorenzo', 'Trooper', 'Thor', 'Cermet', 'quite', 'Marlee',
                'Einstein', 'Rumpole', 'Benny', 'Jarod', 'Iggy', 'Snoop', 'Kyle',
```



'Riley', 'Boomer', 'Noosh', 'Rontu', 'Furzey', 'Daisy', 'Tuck',  
 'Barney', 'Vixen', 'Mimosa', 'Tycho', 'Indie', 'George', 'Bentley',  
 'Maddie', 'Monty', 'Sojourner', 'Odie', 'Stanley', 'Waffles',  
 'Vincent', 'Oliver', 'Eli', 'Tucker', 'Wilson', 'Charlie', 'Bronte',  
 'Poppy', 'Gidget', 'Rhino', 'Willow', 'Eevee', 'Miguel', 'Kuyu',  
 'Lilly', 'Samson', 'Leo', 'Astrid', 'Malcolm', 'Dexter', 'Fiona',  
 'one', 'Sailer', 'Brutus', 'Kona', 'Cupid', 'Pawnd', 'Toby',  
 'Pablo', 'Scooter', 'Nala', 'Wyatt', 'Gabe', 'Sunny', 'Howie',  
 'Anna', 'Bo', 'Bear', 'Autumn', 'Eugene', 'Strudel', 'Binky',  
 'Moose', 'Comet', 'Jack', 'Akumi', 'Olivia', 'Gary', 'Atlas',  
 'Layla', 'Baron', 'Swagger', 'Moe', 'Hunter', 'Phil', 'Kyro',  
 'Ollie', 'Sansa', 'Diogi', 'Severus', 'Loki', 'Bones', 'Mauve',  
 'Doc', 'Peaches', 'Sobe', 'Jeffrey', 'Mister', 'Iroh', 'Cassie',  
 'Pancake', 'Snicku', 'Ruby', 'Brody', 'Bailey', 'Laika', 'Maximus',  
 'Moreton', 'Juno', 'Lily', nan, 'BeBe', 'Remus', 'Levi', 'Mabel',  
 'Duke', 'Brownie', 'Rizzy', 'Frank', 'Tonks', 'Rizzo', 'Pinot',  
 'Dallas', 'Hero', 'Frankie', 'Mairi', 'Loomis', 'Finn', 'Godi',  
 'Dave', 'Earl', 'Cali', 'Sampson', 'Oakley', 'Jay', 'Strider',  
 'Penny', 'Solomon', 'Huck', 'Blue', 'Finley', 'Sprinkles', 'Winnie',  
 'Grey', 'Fizz', 'Klein', 'Tito', 'Rupert', 'Shadoe', 'Tove',  
 'Louie', 'Kota', 'Belle', 'Buddy', 'Shelby', 'Rose', 'Maggie',  
 'Emma', 'Luna', 'Spencer', 'Lilli', 'Leonard', 'Beckham', 'Devón',  
 'Tayzie', 'Max', 'Milo', 'Arnie', 'Calvin', 'Lilah', 'Winston',  
 'Tripp', 'Bookstore', 'Atticus', 'Clark', 'Shiloh', 'Arlen',  
 'Lenox', 'Sugar', 'Abby', 'Stark', 'Maxaroni', 'Aqua', 'Benji',  
 'Bayley', 'Axel', 'Remy', 'Rory', 'Dakota', 'Terry', 'Larry',  
 'Rooney', 'Puff', 'Coleman', 'all', 'Karll', 'Sprout', 'Luther',  
 'Reginald', 'Panda', 'Ebby', 'Link', 'Stephanus', 'Zeus', 'Oscar',  
 'Thumas', 'Piper', 'Olive', 'Beemo', 'Gunner', 'Lacy', 'Watson',  
 'Cecil', 'Karma', 'Smokey', 'Finnegus', 'Cupcake', 'Kathmandu',  
 'Katie', 'Adele', 'Zara', 'Bisquick', 'Yoda', 'Kaia', 'Archie',  
 'Reagan', 'Doug', 'Colin', 'Brian', 'Olivier', 'Chester', 'Dunkin',  
 'Gizmo', 'Lucy', 'Carl', 'Herald', 'Louis', 'Tyrone', 'Baxter',  
 'Oreo', 'Obi', 'Tess', 'Aspen', 'Ozzie', 'Ellie', 'Kawhi', 'by',  
 'Chompsky', 'Sandy', 'Tassy', 'Lenny', 'Laela', 'Raymond',  
 'Pickles', 'Pippin', 'Mia', 'Marley', 'Andy', 'Churlie', 'Jomathan',  
 'Phineas', 'Tilly', 'Darla', 'Zoey', 'Bruno', 'Stuart', 'Jim',  
 'such', 'Derek', 'Jimbo', 'Maisey', 'Kevin', 'Yogi', 'Noah',  
 'Grizzwald', 'Gus', 'Rey', 'Elliot', 'Duddles', 'Steven', 'Beau',  
 'Benedict', 'Venti', 'Goose', 'Nugget', 'Cash', 'Jed', 'Monkey',  
 'Dewey', 'Walter', 'Shikha', 'Neptune', 'Hobbes', 'Burt', 'Jordy',  
 'Milky', 'Sophie', 'Rosie', 'Callie', 'Arya', 'Alice', 'Wiggles',  
 'General', 'Sailor', 'Odin', 'Jerry', 'Georgie', 'Cannon', 'Jarvis',  
 'Brady', 'Margo', 'Sadie', 'Hank', 'Dawn', 'Arlo', 'Mookie',  
 'Meera', 'Ava', 'Ash', 'Tobi', 'Sunshine', 'Lipton', 'not', 'Orion',  
 'Smiley', 'Emanuel', 'Dutch', 'Pete', 'Reggie', 'Alfie', 'Mutt',  
 'Doobert', 'Beebop', 'Alexander', 'Boots', 'Ralphie', 'Pilot',  
 'Ike', 'Mo', 'Sweet', 'Crawford', 'Jimison', 'Duchess', 'Harlso',

'Sundance', 'Luca', 'Flash', 'Jazzy', 'Wafer', 'Tom', 'Florence',  
 'Dido', 'Ken', 'Tebow', 'Chloe', 'Timber', 'Dudley', 'Titan', 'Alf',  
 'Oshie', 'Chubbs', 'Sky', 'Eleanor', 'Rocky', 'Tyr', 'Bauer',  
 'Brandi', 'Mary', 'Halo', 'Augie', 'Craig', 'Pavlov', 'Wallace',  
 'Ito', 'Seamus', 'Stephan', 'Lennon', 'incredibly', 'Major',  
 'Shooter', 'Django', 'Sonny', 'Ronnie', 'Chef', 'Longfellow',  
 'Mack', 'Nimbus', 'Dobby', 'Maude', 'Newt', 'Nida', 'Robin',  
 'Monster', 'Misty', 'Betty', 'Mosby', 'Bruce', 'Happy', 'Ralph',  
 'Stella', 'Butter', 'Lincoln', 'Logan', 'Dale', 'Mattie', 'Stormy',  
 'Balto', 'Deacon', 'Timmy', 'Chipson', 'Combo', 'Dash', 'Hercules',  
 'Mya', 'O', 'Anakin', 'Heinrich', 'Shakespeare', 'Chelsea',  
 'Bungalo', 'Chip', 'Roosevelt', 'Willem', 'Davey', 'Dixie', 'very',  
 'Al', 'Jackson', 'Carbon', 'DonDon', 'Kirby', 'Lou', 'Chevy',  
 'Philbert', 'Rufus', 'Brudge', 'Angel', 'Brat', 'my', 'Gromit',  
 'Aubie', 'Leela', 'Glenn', 'Sephie', 'Bonaparte', 'Albert',  
 'Wishes', 'Theo', 'Rocco', 'Boston', 'Brandonald', 'Corey', 'Gert',  
 'Keith', 'Dex', 'Ace', 'Grizzie', 'Fred', 'Gilbert', 'Meyer', 'Zoe',  
 'Stewie', 'Spanky', 'Jameson', 'Blu', 'Dietrich', 'Divine', 'his',  
 'Cora', 'Huxley', 'Keurig', 'Linus', 'Gustav', 'Percy', 'Harvey',  
 'Blanket', 'actually', 'Geno', 'Beya', 'Kilo', 'Kayla', 'Bell',  
 'Edmund', 'Theodore', 'just', 'Baloo', 'Chase', 'getting', 'Nollie',  
 'Rorie', 'Simba', 'Charles', 'Storkson', 'Chadrick', 'mad',  
 'Kellogg', 'Buckley', 'Livvie', 'Hermione', 'Ralph', 'Aldrick',  
 'unacceptable', 'Crystal', 'Ziva', 'Stefan', 'Pupcasso', 'Flurp',  
 'Enchilada', 'Rueben', 'Cilantro', 'Blitz', 'Bloop', 'Colby',  
 'Lillie', 'Ashleigh', 'Kreggory', 'Sarge', 'Ivar', 'Jangle',  
 'Schnitzel', 'Berkeley', 'Ralphé', 'Charleson', 'Clyde', 'Harnold',  
 'Sid', 'Pippa', 'Otis', 'Carper', 'Bowie', 'Suki', 'Barclay',  
 'Skittle', 'Flávio', 'Jennifur', 'Ozzy', 'Bubbles', 'old',  
 'Bertson', 'Nico', 'Michelangelo', 'Siba', 'Calbert', 'Curtis',  
 'Travis', 'Kanu', 'Lance', 'Opie', 'Stubert', 'Kane', 'Chuckles',  
 'Staniel', 'Sora', 'infuriating', 'Tater', 'Olaf', 'Vince', 'Billy',  
 'Walker', 'Rodney', 'Klevin', 'Malikai', 'Bobbie', 'River',  
 'Jebberson', 'Remington', 'Farfle', 'Jiminus', 'Harper', 'Clarkus',  
 'Kara', 'Ambrose', 'Jimothy', 'Bode', 'Terrenth', 'Reese',  
 'Chesterson', 'Lucia', 'Ralphson', 'Socks', 'Rambo', 'Rudy', 'Fiji',  
 'Rilo', 'Bilbo', 'Coopson', 'Millie', 'Chet', 'Croun', 'Daniel',  
 'Murphy', 'Dotsy', 'Eazy', 'Coops', 'Fillup', 'Miley', 'Yukon',  
 'CeCe', 'Cuddles', 'Claude', 'Jessiga', 'Carter', 'Ole', 'Pherb',  
 'Blipson', 'Reptar', 'Trevith', 'Berb', 'Bob', 'Grady', 'Kobe',  
 'Freddery', 'Bodie', 'Wally', 'Tupawc', 'Amber', 'Herschel',  
 'Edgar', 'Teddy', 'Kingsley', 'Brockly', 'Richie', 'Molly',  
 'Vinscent', 'Cedrick', 'Hazel', 'Lolo', 'Eriq', 'Phred', 'the',  
 'Oddie', 'Maxwell', 'Geoff', 'Covach', 'Durg', 'Fynn', 'Ricky',  
 'Lucky', 'Ferg', 'Trip', 'Clarence', 'Hamrick', 'Brad', 'Pubert',  
 'Frönq', 'Derby', 'Lizzie', 'Ember', 'Blakely', 'Opal', 'Marq',  
 'Kramer', 'Barry', 'Gordon', 'Mona', 'Horace', 'Crimson', 'Birf',  
 'Lorelei', 'Marty', 'Brooks', 'Petrick', 'Hubertson', 'Gerbald',

```
'Bruiser', 'Perry', 'Jeph', 'Tino', 'Kulet', 'Sweets', 'Lupe',
'Tiger', 'Griffin', 'Banjo', 'Brandy', 'Lulu', 'Taco', 'Joey',
'Patrick', 'Kreg', 'Todo', 'Ulysses', 'Toffee', 'Apollo', 'Carly',
'Asher', 'Glacier', 'Champ', 'Griswold', 'Cheesy', 'Moofasa',
'Hector', 'Goliath', 'Emmie', 'Willie', 'Rinna', 'Mike', 'William',
'Dwight', 'Evy', 'Hurley', 'Rubio', 'officially', 'Rascal', 'Linda',
'Tug', 'Tango', 'Grizz', 'Jerome', 'Crumpet', 'Jessifer', 'Izzy',
'Ralph', 'Humphrey', 'Chuq', 'Tyrus', 'Karl', 'Godzilla', 'Vinnie',
'Kenneth', 'Herm', 'Bert', 'Striker', 'Donny', 'Pepper', 'Bernie',
'Buddah', 'Wylie', 'Arnold', 'Zuzu', 'Mollie', 'Tedders',
'Superpup', 'Rufio', 'Jeb', 'Rodman', 'Jonah', 'Chesney', 'life',
'Kenny', 'Bobbay', 'Mitch', 'Kaiya', 'Acro', 'Aiden', 'Obie', 'Dot',
'Shnuggles', 'Kendall', 'Kip', 'Jeffri', 'Steve', 'Eve', 'Mac',
'Fletcher', 'Kenzie', 'Pumpkin', 'Schnozz', 'Gustaf', 'Cheryl',
'Ed', 'Leonidas', 'Norman', 'Caryl', 'Scott', 'Taz', 'Darby',
'Jackie', 'light', 'Jazz', 'Franq', 'Rolf', 'Snickers', 'Ridley',
'Cal', 'Bradley', 'Bubba', 'Tuco', 'Patch', 'Mojo', 'Batdog',
'Dylan', 'space', 'Mark', 'Jacob', 'JD', 'Alejandro', 'Scruffers',
'Hammond', 'Pip', 'Julius', 'Penelope', 'Tanner', 'Sparky', 'Holly',
'Jett', 'Amy', 'Sage', 'Mason', 'Trigger', 'Antony', 'Creg',
'Traviss', 'Gin', 'Jeffrie', 'Danny', 'Ester', 'Pluto', 'Bloo',
'Edd', 'Paull', 'Sam', 'Willy', 'Spork', 'Herb', 'Damon', 'Peanut',
'Nigel', 'Cherokee', 'Butters', 'Hemry', 'Sandra', 'Fabio', 'Liam',
'Ben', 'Raphael', 'Julio', 'Andru', 'Alphred', 'Kloey', 'Shawwn',
'Skye', 'Kollin', 'Alfredo', 'Ronduh', 'Billl', 'Saydee', 'Dug',
'Sully', 'Kirk', 'Ralf', 'Clarq', 'Jaspers', 'Samsom', 'Terrance',
'Harrison', 'Chaz', 'Jeremy', 'Jaycob', 'Leroi', 'Lambeau',
'Ruffles', 'Amélie', 'Bobb', 'Banditt', 'Kevon', 'Winifred', 'Hanz',
'Berta', 'Timofy', 'Maks', 'Kallie', 'Marvin', 'Spark', 'Gördón',
'Chuk', 'Jo', 'DayZ', 'Jareld', 'Torque', 'Ron', 'Skittles',
'Alfonso', 'Jiminy', 'Cleopatra', 'Erik', 'Stu', 'Shaggy',
'Filup', 'Kial', 'Klint', 'Naphaniel', 'Dook', 'Hall', 'Philippe',
'Kohl', 'Biden', 'Fwed', 'Genevieve', 'Joshwa', 'Timison', 'Daryl',
'Bradlay', 'Pipsy', 'Clybe', 'Keet', 'Carll', 'Pepe', 'Jockson',
'Octaviath', 'Lugan', 'Johm', 'Christoper'], dtype=object)
```

## Define

- df\_2 dataset : img\_num column == consistency .

## Code

```
In [54]: ## drop img_num column
df2_clean = df2_clean.drop(['img_num'] , axis = 1)
```

## Test

```
In [55]: df2_clean
```

Out [55]:	tweet_id	jpg_url \
0	666020888022790149	<a href="https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg">https://pbs.twimg.com/media/CT4udn0WwAA0aMy.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/CT5w9gUW4AAsBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAsBNN.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/CT56LSZW0AA1Jj2.jpg">https://pbs.twimg.com/media/CT56LSZW0AA1Jj2.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/CT9Vn7PW0AA_ZCM.jpg">https://pbs.twimg.com/media/CT9Vn7PW0AA_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-NvwmW4AAUGZ.jpg">https://pbs.twimg.com/media/CT-NvwmW4AAUGZ.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/DE8yicJW0AAAABJ.jpg">https://pbs.twimg.com/media/DE8yicJW0AAAABJ.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-UQAACu20.jpg">https://pbs.twimg.com/media/DFTH_0-UQAACu20.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/DF6hr6BUMAAZgT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAZgT.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>

	p1	p1_conf	p1_dog	p2 \
0	Welsh_springer_spaniel	0.465074	True	collie
1	redbone	0.506826	True	miniature_pinscher
2	German_shepherd	0.596461	True	malinois
3	Rhodesian_ridgeback	0.408143	True	redbone
4	miniature_pinscher	0.560311	True	Rottweiler
5	Bernese_mountain_dog	0.651137	True	English_springer
6	box_turtle	0.933012	False	mud_turtle
7	chow	0.692517	True	Tibetan_mastiff
8	shopping_cart	0.962465	False	shopping_basket
9	miniature_poodle	0.201493	True	komondor
10	golden_retriever	0.775930	True	Tibetan_mastiff
11	Gordon_setter	0.503672	True	Yorkshire_terrier
12	Walker_hound	0.260857	True	English_foxhound
13	pug	0.489814	True	bull_mastiff
14	bloodhound	0.195217	True	German_shepherd
15	Lhasa	0.582330	True	Shih-Tzu
16	English_setter	0.298617	True	Newfoundland
17	hen	0.965932	False	cock
18	desktop_computer	0.086502	False	desk
19	Italian_greyhound	0.176053	True	toy_terrier
20	Maltese_dog	0.857531	True	toy_poodle
21	three-toed_sloth	0.914671	False	otter
22	ox	0.416669	False	Newfoundland
23	golden_retriever	0.858744	True	Chesapeake_Bay_retriever
24	malamute	0.336874	True	Siberian_husky
25	guinea_pig	0.996496	False	skunk
26	soft-coated_wheaten_terrier	0.326467	True	Afghan_hound
27	Chihuahua	0.978108	True	toy_terrier
28	black-and-tan_coonhound	0.529139	True	bloodhound
29	coho	0.404640	False	barracouta
...	...	...	...	...
2045	French_bulldog	0.999201	True	Chihuahua

2046	convertible	0.738995	False	sports_car
2047	kuvasz	0.309706	True	Great-Pyrenees
2048	Chihuahua	0.793469	True	toy_terrier
2049	Samoyed	0.733942	True	Eskimo_dog
2050	Mexican_hairless	0.330741	True	sea_lion
2051	Pembroke	0.809197	True	Rhodesian_ridgeback
2052	limousine	0.130432	False	tow_truck
2053	basset	0.821664	True	redbone
2054	French_bulldog	0.995026	True	pug
2055	Pembroke	0.809197	True	Rhodesian_ridgeback
2056	Siberian_husky	0.700377	True	Eskimo_dog
2057	golden_retriever	0.469760	True	Labrador_retriever
2058	golden_retriever	0.714719	True	Tibetan_mastiff
2059	whippet	0.626152	True	borzoi
2060	golden_retriever	0.953442	True	Labrador_retriever
2061	French_bulldog	0.991650	True	boxer
2062	Pembroke	0.966327	True	Cardigan
2063	French_bulldog	0.377417	True	Labrador_retriever
2064	Samoyed	0.957979	True	Pomeranian
2065	Pembroke	0.511319	True	Cardigan
2066	Irish_terrier	0.487574	True	Irish_setter
2067	Pomeranian	0.566142	True	Eskimo_dog
2068	Appenzeller	0.341703	True	Border_collie
2069	Chesapeake_Bay_retriever	0.425595	True	Irish_terrier
2070	basset	0.555712	True	English_springer
2071	paper_towel	0.170278	False	Labrador_retriever
2072	Chihuahua	0.716012	True	malamute
2073	Chihuahua	0.323581	True	Pekinese
2074	orange	0.097049	False	bagel

	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	0.263788	True	Greater_Swiss_Mountain_dog	0.016199	True	
6	0.045885	False	terrapin	0.017885	False	
7	0.058279	True	fur_coat	0.054449	False	
8	0.014594	False	golden_retriever	0.007959	True	
9	0.192305	True	soft-coated_wheaten_terrier	0.082086	True	
10	0.093718	True	Labrador_retriever	0.072427	True	
11	0.174201	True	Pekinese	0.109454	True	
12	0.175382	True	Ibizan_hound	0.097471	True	
13	0.404722	True	French_bulldog	0.048960	True	
14	0.078260	True	malinois	0.075628	True	
15	0.166192	True	Dandie_Dinmont	0.089688	True	
16	0.149842	True	borzoi	0.133649	True	

17	0.033919	False	partridge	0.000052	False
18	0.085547	False	bookcase	0.079480	False
19	0.111884	True	basenji	0.111152	True
20	0.063064	True	miniature_poodle	0.025581	True
21	0.015250	False	great_grey_owl	0.013207	False
22	0.278407	True	groenendael	0.102643	True
23	0.054787	True	Labrador_retriever	0.014241	True
24	0.147655	True	Eskimo_dog	0.093412	True
25	0.002402	False	hamster	0.000461	False
26	0.259551	True	briard	0.206803	True
27	0.009397	True	papillon	0.004577	True
28	0.244220	True	flat-coated_retriever	0.173810	True
29	0.271485	False	gar	0.189945	False
...	...	...	...	...	...
2045	0.000361	True	Boston_bull	0.000076	True
2046	0.139952	False	car_wheel	0.044173	False
2047	0.186136	True	Dandie_Dinmont	0.086346	True
2048	0.143528	True	can_opener	0.032253	False
2049	0.035029	True	Staffordshire_bullterrier	0.029705	True
2050	0.275645	False	Weimaraner	0.134203	True
2051	0.054950	True	beagle	0.038915	True
2052	0.029175	False	shopping_cart	0.026321	False
2053	0.087582	True	Weimaraner	0.026236	True
2054	0.000932	True	bull_mastiff	0.000903	True
2055	0.054950	True	beagle	0.038915	True
2056	0.166511	True	malamute	0.111411	True
2057	0.184172	True	English_setter	0.073482	True
2058	0.120184	True	Labrador_retriever	0.105506	True
2059	0.194742	True	Saluki	0.027351	True
2060	0.013834	True	redbone	0.007958	True
2061	0.002129	True	Staffordshire_bullterrier	0.001498	True
2062	0.027356	True	basenji	0.004633	True
2063	0.151317	True	muzzle	0.082981	False
2064	0.013884	True	chow	0.008167	True
2065	0.451038	True	Chihuahua	0.029248	True
2066	0.193054	True	Chesapeake_Bay_retriever	0.118184	True
2067	0.178406	True	Pembroke	0.076507	True
2068	0.199287	True	ice_lolly	0.193548	False
2069	0.116317	True	Indian_elephant	0.076902	False
2070	0.225770	True	German_short-haired_pointer	0.175219	True
2071	0.168086	True	spatula	0.040836	False
2072	0.078253	True	kelpie	0.031379	True
2073	0.090647	True	papillon	0.068957	True
2074	0.085851	False	banana	0.076110	False

[2075 rows x 11 columns]

In [56]: df2\_clean.info()

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2075 entries, 0 to 2074
Data columns (total 11 columns):
tweet_id      2075 non-null int64
jpg_url       2075 non-null object
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(1), object(4)
memory usage: 135.8+ KB

```

## Define

- df\_2 dataset : df\_2 should contain only the highest prediction confident of breed and the breed name instead of 3 predictions.

## Code

In [57]: df2\_clean

```

Out[57]:
      tweet_id      jpg_url \
0    666020888022790149  https://pbs.twimg.com/media/CT4udn0WwAA0aMy.jpg
1    666029285002620928  https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg
2    666033412701032449  https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg
3    666044226329800704  https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg
4    666049248165822465  https://pbs.twimg.com/media/CT5IQmsXIAAKY4A.jpg
5    666050758794694657  https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg
6    666051853826850816  https://pbs.twimg.com/media/CT5KoJ1WoAAJash.jpg
7    666055525042405380  https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg
8    666057090499244032  https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg
9    666058600524156928  https://pbs.twimg.com/media/CT5Qw94XAAA-2dP.jpg
10   666063827256086533  https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg
11   666071193221509120  https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg
12   666073100786774016  https://pbs.twimg.com/media/CT5d9DZXAAALcwe.jpg
13   666082916733198337  https://pbs.twimg.com/media/CT5m4VGWEAAKc8.jpg
14   666094000022159362  https://pbs.twimg.com/media/CT5w9gUW4AAAsBNN.jpg
15   666099513787052032  https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg
16   666102155909144576  https://pbs.twimg.com/media/CT54YGiWUAEZnoK.jpg
17   666104133288665088  https://pbs.twimg.com/media/CT56LSZWAAALJj2.jpg
18   666268910803644416  https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg
19   666273097616637952  https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg

```



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-NvwmW4AAAugGZ.jpg">https://pbs.twimg.com/media/CT-NvwmW4AAAugGZ.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/DE8yicJW0AAAABJ.jpg">https://pbs.twimg.com/media/DE8yicJW0AAAABJ.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-UQAACu20.jpg">https://pbs.twimg.com/media/DFTH_0-UQAACu20.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/DF1e0mZXUAAALUc.jpg">https://pbs.twimg.com/media/DF1e0mZXUAAALUc.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/DF6hr6BUMAAZgT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAZgT.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>

		p1	p1_conf	p1_dog	p2 \
0	Welsh_springer_spaniel	0.465074	True		collie
1	redbone	0.506826	True		miniature_pinscher
2	German_shepherd	0.596461	True		malinois
3	Rhodesian_ridgeback	0.408143	True		redbone
4	miniature_pinscher	0.560311	True		Rottweiler

5	Bernese_mountain_dog	0.651137	True	English_springer
6	box_turtle	0.933012	False	mud_turtle
7	chow	0.692517	True	Tibetan_mastiff
8	shopping_cart	0.962465	False	shopping_basket
9	miniature_poodle	0.201493	True	komondor
10	golden_retriever	0.775930	True	Tibetan_mastiff
11	Gordon_setter	0.503672	True	Yorkshire_terrier
12	Walker_hound	0.260857	True	English_foxhound
13	pug	0.489814	True	bull_mastiff
14	bloodhound	0.195217	True	German_shepherd
15	Lhasa	0.582330	True	Shih-Tzu
16	English_setter	0.298617	True	Newfoundland
17	hen	0.965932	False	cock
18	desktop_computer	0.086502	False	desk
19	Italian_greyhound	0.176053	True	toy_terrier
20	Maltese_dog	0.857531	True	toy_poodle
21	three-toed_sloth	0.914671	False	otter
22	ox	0.416669	False	Newfoundland
23	golden_retriever	0.858744	True	Chesapeake_Bay_retriever
24	malamute	0.336874	True	Siberian_husky
25	guinea_pig	0.996496	False	skunk
26	soft-coated_wheaten_terrier	0.326467	True	Afghan_hound
27	Chihuahua	0.978108	True	toy_terrier
28	black-and-tan_coonhound	0.529139	True	bloodhound
29	coho	0.404640	False	barracouta
...	...	...	...	...
2045	French_bulldog	0.999201	True	Chihuahua
2046	convertible	0.738995	False	sports_car
2047	kuvasz	0.309706	True	Great_Pyrenees
2048	Chihuahua	0.793469	True	toy_terrier
2049	Samoyed	0.733942	True	Eskimo_dog
2050	Mexican_hairless	0.330741	True	sea_lion
2051	Pembroke	0.809197	True	Rhodesian_ridgeback
2052	limousine	0.130432	False	tow_truck
2053	basset	0.821664	True	redbone
2054	French_bulldog	0.995026	True	pug
2055	Pembroke	0.809197	True	Rhodesian_ridgeback
2056	Siberian_husky	0.700377	True	Eskimo_dog
2057	golden_retriever	0.469760	True	Labrador_retriever
2058	golden_retriever	0.714719	True	Tibetan_mastiff
2059	whippet	0.626152	True	borzoi
2060	golden_retriever	0.953442	True	Labrador_retriever
2061	French_bulldog	0.991650	True	boxer
2062	Pembroke	0.966327	True	Cardigan
2063	French_bulldog	0.377417	True	Labrador_retriever
2064	Samoyed	0.957979	True	Pomeranian
2065	Pembroke	0.511319	True	Cardigan
2066	Irish_terrier	0.487574	True	Irish_setter

2067	Pomeranian	0.566142	True	Eskimo_dog
2068	Appenzeller	0.341703	True	Border_collie
2069	Chesapeake_Bay_retriever	0.425595	True	Irish_terrier
2070	basset	0.555712	True	English_springer
2071	paper_towel	0.170278	False	Labrador_retriever
2072	Chihuahua	0.716012	True	malamute
2073	Chihuahua	0.323581	True	Pekinese
2074	orange	0.097049	False	bagel

	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	0.263788	True	Greater_Swiss_Mountain_dog	0.016199	True	
6	0.045885	False	terrapin	0.017885	False	
7	0.058279	True	fur_coat	0.054449	False	
8	0.014594	False	golden_retriever	0.007959	True	
9	0.192305	True	soft-coated_wheaten_terrier	0.082086	True	
10	0.093718	True	Labrador_retriever	0.072427	True	
11	0.174201	True	Pekinese	0.109454	True	
12	0.175382	True	Ibizan_hound	0.097471	True	
13	0.404722	True	French_bulldog	0.048960	True	
14	0.078260	True	malinois	0.075628	True	
15	0.166192	True	Dandie_Dinmont	0.089688	True	
16	0.149842	True	borzoi	0.133649	True	
17	0.033919	False	partridge	0.000052	False	
18	0.085547	False	bookcase	0.079480	False	
19	0.111884	True	basenji	0.111152	True	
20	0.063064	True	miniature_poodle	0.025581	True	
21	0.015250	False	great_grey_owl	0.013207	False	
22	0.278407	True	groenendael	0.102643	True	
23	0.054787	True	Labrador_retriever	0.014241	True	
24	0.147655	True	Eskimo_dog	0.093412	True	
25	0.002402	False	hamster	0.000461	False	
26	0.259551	True	briard	0.206803	True	
27	0.009397	True	papillon	0.004577	True	
28	0.244220	True	flat-coated_retriever	0.173810	True	
29	0.271485	False	gar	0.189945	False	
...	...	...	...	...	...	...
2045	0.000361	True	Boston_bull	0.000076	True	
2046	0.139952	False	car_wheel	0.044173	False	
2047	0.186136	True	Dandie_Dinmont	0.086346	True	
2048	0.143528	True	can_opener	0.032253	False	
2049	0.035029	True	Staffordshire_bullterrier	0.029705	True	
2050	0.275645	False	Weimaraner	0.134203	True	
2051	0.054950	True	beagle	0.038915	True	

2052	0.029175	False	shopping_cart	0.026321	False
2053	0.087582	True	Weimaraner	0.026236	True
2054	0.000932	True	bull_mastiff	0.000903	True
2055	0.054950	True	beagle	0.038915	True
2056	0.166511	True	malamute	0.111411	True
2057	0.184172	True	English_setter	0.073482	True
2058	0.120184	True	Labrador_retriever	0.105506	True
2059	0.194742	True	Saluki	0.027351	True
2060	0.013834	True	redbone	0.007958	True
2061	0.002129	True	Staffordshire_bullterrier	0.001498	True
2062	0.027356	True	basenji	0.004633	True
2063	0.151317	True	muzzle	0.082981	False
2064	0.013884	True	chow	0.008167	True
2065	0.451038	True	Chihuahua	0.029248	True
2066	0.193054	True	Chesapeake_Bay_retriever	0.118184	True
2067	0.178406	True	Pembroke	0.076507	True
2068	0.199287	True	ice_lolly	0.193548	False
2069	0.116317	True	Indian_elephant	0.076902	False
2070	0.225770	True	German_short-haired_pointer	0.175219	True
2071	0.168086	True	spatula	0.040836	False
2072	0.078253	True	kelpie	0.031379	True
2073	0.090647	True	papillon	0.068957	True
2074	0.085851	False	banana	0.076110	False

[2075 rows x 11 columns]

In [58]: *## choose the maximum p\_conf which will be the breed , use conditional method.*

```
df2_clean["pic_conf"] = df2_clean[["p1_conf", "p2_conf", "p3_conf"]].max(axis=1)
```

```
conditions = [(df2_clean["pic_conf"] == df2_clean['p1_conf'] ) , (df2_clean["pic_conf"]
                    (df2_clean["pic_conf"] == df2_clean['p2_conf'] ) )]
```

```
choices = [df2_clean['p1'], df2_clean['p2'], df2_clean['p3']]
```

```
df2_clean["breed"] = np.select(conditions, choices)
```

*## drop p and conf columns not required.*

```
df2_clean = df2_clean.drop(['p1', 'p1_conf', 'p1_dog', 'p2', 'p2_conf', 'p2_dog', 'p3', 'p3_co
```

## Test

In [59]: df2\_clean

```
Out[59]:
```

	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/CT5w9gUW4AAAsBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAAsBNN.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/CT56LSZWoAA1Jj2.jpg">https://pbs.twimg.com/media/CT56LSZWoAA1Jj2.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/CT9Vn7PW0AA_ZCM.jpg">https://pbs.twimg.com/media/CT9Vn7PW0AA_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/CT9vZEYUAA1Z05.jpg">https://pbs.twimg.com/media/CT9vZEYUAA1Z05.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-NvwmW4AAUGZ.jpg">https://pbs.twimg.com/media/CT-NvwmW4AAUGZ.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/DE5Se8FXcAAAJFx4.jpg">https://pbs.twimg.com/media/DE5Se8FXcAAAJFx4.jpg</a>
2048	886983233522544640	<a href="https://pbs.twimg.com/media/DE8yicJW0AAAABJ.jpg">https://pbs.twimg.com/media/DE8yicJW0AAAABJ.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/DFHQBbXgAEqY7t.jpg">https://pbs.twimg.com/media/DFHQBbXgAEqY7t.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/DFrEyVuWAAA03t9.jpg">https://pbs.twimg.com/media/DFrEyVuWAAA03t9.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/DF1e0mZXUAAULUc.jpg">https://pbs.twimg.com/media/DF1e0mZXUAAULUc.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/DF6hr6BUMAAZgT.jpg">https://pbs.twimg.com/media/DF6hr6BUMAAZgT.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>

	pic_conf	breed
0	0.465074	Welsh_springer_spaniel
1	0.506826	redbone
2	0.596461	German_shepherd
3	0.408143	Rhodesian_ridgeback
4	0.560311	miniature_pinscher
5	0.651137	Bernese_mountain_dog
6	0.933012	box_turtle
7	0.692517	chow
8	0.962465	shopping_cart
9	0.201493	miniature_poodle
10	0.775930	golden_retriever
11	0.503672	Gordon_setter
12	0.260857	Walker_hound
13	0.489814	pug
14	0.195217	bloodhound
15	0.582330	Lhasa
16	0.298617	English_setter
17	0.965932	hen
18	0.086502	desktop_computer
19	0.176053	Italian_greyhound
20	0.857531	Maltese_dog
21	0.914671	three-toed_sloth
22	0.416669	ox
23	0.858744	golden_retriever
24	0.336874	malamute
25	0.996496	guinea_pig
26	0.326467	soft-coated_wheaten_terrier
27	0.978108	Chihuahua
28	0.529139	black-and-tan_coonhound
29	0.404640	coho
...	...	...
2045	0.999201	French_bulldog
2046	0.738995	convertible
2047	0.309706	kuvasz
2048	0.793469	Chihuahua
2049	0.733942	Samoyed

2050	0.330741	Mexican_hairless
2051	0.809197	Pembroke
2052	0.130432	limousine
2053	0.821664	basset
2054	0.995026	French_bulldog
2055	0.809197	Pembroke
2056	0.700377	Siberian_husky
2057	0.469760	golden_retriever
2058	0.714719	golden_retriever
2059	0.626152	whippet
2060	0.953442	golden_retriever
2061	0.991650	French_bulldog
2062	0.966327	Pembroke
2063	0.377417	French_bulldog
2064	0.957979	Samoyed
2065	0.511319	Pembroke
2066	0.487574	Irish_terrier
2067	0.566142	Pomeranian
2068	0.341703	Appenzeller
2069	0.425595	Chesapeake_Bay_retriever
2070	0.555712	basset
2071	0.170278	paper_towel
2072	0.716012	Chihuahua
2073	0.323581	Chihuahua
2074	0.097049	orange

[2075 rows x 4 columns]

## Define

- df\_1 , df\_2 ,df\_3 datasets : all should be in one dataset.

## Code

```
In [60]: sub_master = pd.merge(df1_clean, df2_clean , how='inner' , on=["tweet_id"])
twitter_master = pd.merge(sub_master, df3_clean , how='inner' , on=["tweet_id"])
twitter_master.reset_index(drop = True)
```

```
Out[60]:
```

	tweet_id	timestamp	source \
0	891327558926688256	2017-07-29 16:00:24	twitter for iphone
1	890971913173991426	2017-07-28 16:27:12	twitter for iphone
2	890729181411237888	2017-07-28 00:22:40	twitter for iphone
3	890006608113172480	2017-07-26 00:31:25	twitter for iphone
4	889638837579907072	2017-07-25 00:10:02	twitter for iphone
5	888804989199671297	2017-07-22 16:56:37	twitter for iphone
6	888554962724278272	2017-07-22 00:23:06	twitter for iphone
7	888078434458587136	2017-07-20 16:49:33	twitter for iphone
8	887473957103951883	2017-07-19 00:47:34	twitter for iphone

9	886983233522544640	2017-07-17	16:17:36	twitter for iphone
10	886736880519319552	2017-07-16	23:58:41	twitter for iphone
11	886366144734445568	2017-07-15	23:25:31	twitter for iphone
12	885167619883638784	2017-07-12	16:03:00	twitter for iphone
13	884876753390489601	2017-07-11	20:47:12	twitter for iphone
14	883117836046086144	2017-07-07	00:17:54	twitter for iphone
15	882992080364220416	2017-07-06	15:58:11	twitter for iphone
16	882268110199369728	2017-07-04	16:01:23	twitter for iphone
17	882045870035918850	2017-07-04	01:18:17	twitter for iphone
18	880465832366813184	2017-06-29	16:39:47	twitter for iphone
19	880221127280381952	2017-06-29	00:27:25	twitter for iphone
20	879862464715927552	2017-06-28	00:42:13	twitter for iphone
21	878776093423087618	2017-06-25	00:45:22	twitter for iphone
22	878281511006478336	2017-06-23	16:00:04	twitter for iphone
23	878057613040115712	2017-06-23	01:10:23	twitter for iphone
24	877736472329191424	2017-06-22	03:54:17	twitter for iphone
25	877316821321428993	2017-06-21	00:06:44	twitter for iphone
26	877201837425926144	2017-06-20	16:29:50	twitter for iphone
27	876838120628539392	2017-06-19	16:24:33	twitter for iphone
28	875021211251597312	2017-06-14	16:04:48	twitter for iphone
29	874012996292530176	2017-06-11	21:18:31	twitter for iphone
...	...	...	...	...
1841	666437273139982337	2015-11-17	02:06:42	twitter for iphone
1842	666435652385423360	2015-11-17	02:00:15	twitter for iphone
1843	666430724426358785	2015-11-17	01:40:41	twitter for iphone
1844	666428276349472768	2015-11-17	01:30:57	twitter for iphone
1845	666421158376562688	2015-11-17	01:02:40	twitter for iphone
1846	666418789513326592	2015-11-17	00:53:15	twitter for iphone
1847	666407126856765440	2015-11-17	00:06:54	twitter for iphone
1848	666396247373291520	2015-11-16	23:23:41	twitter for iphone
1849	666373753744588802	2015-11-16	21:54:18	twitter for iphone
1850	666362758909284353	2015-11-16	21:10:36	twitter for iphone
1851	666353288456101888	2015-11-16	20:32:58	twitter for iphone
1852	666345417576210432	2015-11-16	20:01:42	twitter for iphone
1853	666337882303524864	2015-11-16	19:31:45	twitter for iphone
1854	666273097616637952	2015-11-16	15:14:19	twitter for iphone
1855	666268910803644416	2015-11-16	14:57:41	twitter for iphone
1856	666102155909144576	2015-11-16	03:55:04	twitter for iphone
1857	666099513787052032	2015-11-16	03:44:34	twitter for iphone
1858	666094000022159362	2015-11-16	03:22:39	twitter for iphone
1859	666082916733198337	2015-11-16	02:38:37	twitter for iphone
1860	666073100786774016	2015-11-16	01:59:36	twitter for iphone
1861	666071193221509120	2015-11-16	01:52:02	twitter for iphone
1862	666063827256086533	2015-11-16	01:22:45	twitter for iphone
1863	666058600524156928	2015-11-16	01:01:59	twitter for iphone
1864	666057090499244032	2015-11-16	00:55:59	twitter for iphone
1865	666055525042405380	2015-11-16	00:49:46	twitter for iphone
1866	666050758794694657	2015-11-16	00:30:50	twitter for iphone



1867 666044226329800704 2015-11-16 00:04:52 twitter for iphone  
 1868 666033412701032449 2015-11-15 23:21:54 twitter for iphone  
 1869 666029285002620928 2015-11-15 23:05:30 twitter for iphone  
 1870 666020888022790149 2015-11-15 22:32:08 twitter for iphone

text \

0 This is Franklin. He would like you to stop ca...  
 1 Meet Jax. He enjoys ice cream so much he gets ...  
 2 When you watch your owner call another dog a g...  
 3 This is Koda. He is a South Australian decksha...  
 4 This is Ted. He does his best. Sometimes that'...  
 5 This is Zeke. He has a new stick. Very proud o...  
 6 This is Ralphus. He's powering up. Attempting ...  
 7 This is Gerald. He was just told he didn't get...  
 8 This is Canela. She attempted some fancy porch...  
 9 This is Maya. She's very shy. Rarely leaves he...  
 10 This is Mingus. He's a wonderful father to his...  
 11 This is Roscoe. Another pupper fallen victim t...  
 12 Here we have a corgi undercover as a malamute...  
 13 This is Lola. It's her first time outside. Mus...  
 14 Please only send dogs. We don't rate mechanics...  
 15 This is Rusty. He wasn't ready for the first p...  
 16 This is Alf. You're witnessing his first wate...  
 17 This is Koko. Her owner, inspired by Barney, r...  
 18 This is Bella. She had her first beach experie...  
 19 Meet Jesse. He's a Fetty Woof. His tongue ejec...  
 20 This is Romeo. He would like to do an entrance...  
 21 This is Snoopy. He's a proud #PrideMonthPuppo...  
 22 Meet Shadow. In an attempt to reach maximum zo...  
 23 This is Emmy. She was adopted today. Massive r...  
 24 This is Aja. She was just told she's a good do...  
 25 Meet Dante. At first he wasn't a fan of his ne...  
 26 This is Nelly. He graduated with his dogtorate...  
 27 This is Ginger. She's having a ruff Monday. To...  
 28 Guys please stop sending pictures without any ...  
 29 This is Sebastian. He can't see all the colors...  
 ...  
 1841 Here we see a lone northeastern Cumberbatch. H...  
 1842 "Can you behave? You're ruining my wedding day...  
 1843 Oh boy what a pup! Sunglasses take this one to...  
 1844 Here we have an Austrian Pulitzer. Collectors ...  
 1845 \*internally screaming\* 12/10  
 1846 This is Walter. He is an Alaskan Terrapin. Lov...  
 1847 This is a southern Vesuvius bumblegruff. Can d...  
 1848 Oh goodness. A super rare northeast Qdoba kang...  
 1849 Those are sunglasses and a jean jacket. 11/10 ...  
 1850 Unique dog here. Very small. Lives in containe...  
 1851 Here we have a mixed Asiago from the Galápagos...

1852 Look at this jokester thinking seat belt laws ...  
 1853 This is an extremely rare horned Parthenon. No...  
 1854 Can take selfies 11/10  
 1855 Very concerned about fellow dog trapped in com...  
 1856 Oh my. Here you are seeing an Adobe Setter giv...  
 1857 Can stand on stump for what seems like a while...  
 1858 This appears to be a Mongolian Presbyterian mi...  
 1859 Here we have a well-established sunblockerspan...  
 1860 Let's hope this flight isn't Malaysian (lol). ...  
 1861 Here we have a northern speckled Rhododendron...  
 1862 This is the happiest dog you will ever see. Ve...  
 1863 Here is the Rand Paul of retrievers folks! He'...  
 1864 My oh my. This is a rare blond Canadian terrie...  
 1865 Here is a Siberian heavily armored polar bear ...  
 1866 This is a truly beautiful English Wilson Staff...  
 1867 This is a purebred Piers Morgan. Loves to Netf...  
 1868 Here is a very happy pup. Big fan of well-main...  
 1869 This is a western brown Mitsubishi terrier. Up...  
 1870 Here we have a Japanese Irish Setter. Lost eye...

	expanded_urls	rating_numerator \
0	<a href="https://twitter.com/dog_rates/status/891327558...">https://twitter.com/dog_rates/status/891327558...</a>	12.0
1	<a href="https://twitter.com/dog_rates/status/890971913...">https://twitter.com/dog_rates/status/890971913...</a>	13.0
2	<a href="https://twitter.com/dog_rates/status/890729181...">https://twitter.com/dog_rates/status/890729181...</a>	13.0
3	<a href="https://twitter.com/dog_rates/status/890006608...">https://twitter.com/dog_rates/status/890006608...</a>	13.0
4	<a href="https://twitter.com/dog_rates/status/889638837...">https://twitter.com/dog_rates/status/889638837...</a>	12.0
5	<a href="https://twitter.com/dog_rates/status/888804989...">https://twitter.com/dog_rates/status/888804989...</a>	13.0
6	<a href="https://twitter.com/dog_rates/status/888554962...">https://twitter.com/dog_rates/status/888554962...</a>	13.0
7	<a href="https://twitter.com/dog_rates/status/888078434...">https://twitter.com/dog_rates/status/888078434...</a>	12.0
8	<a href="https://twitter.com/dog_rates/status/887473957...">https://twitter.com/dog_rates/status/887473957...</a>	13.0
9	<a href="https://twitter.com/dog_rates/status/886983233...">https://twitter.com/dog_rates/status/886983233...</a>	13.0
10	<a href="https://twitter.com/dog_rates/status/886736880...">https://twitter.com/dog_rates/status/886736880...</a>	13.0
11	<a href="https://twitter.com/dog_rates/status/886366144...">https://twitter.com/dog_rates/status/886366144...</a>	12.0
12	<a href="https://twitter.com/dog_rates/status/885167619...">https://twitter.com/dog_rates/status/885167619...</a>	13.0
13	<a href="https://twitter.com/dog_rates/status/884876753...">https://twitter.com/dog_rates/status/884876753...</a>	13.0
14	<a href="https://twitter.com/dog_rates/status/883117836...">https://twitter.com/dog_rates/status/883117836...</a>	13.0
15	<a href="https://twitter.com/dog_rates/status/882992080...">https://twitter.com/dog_rates/status/882992080...</a>	13.0
16	<a href="https://twitter.com/dog_rates/status/882268110...">https://twitter.com/dog_rates/status/882268110...</a>	13.0
17	<a href="https://twitter.com/dog_rates/status/882045870...">https://twitter.com/dog_rates/status/882045870...</a>	13.0
18	<a href="https://twitter.com/dog_rates/status/880465832...">https://twitter.com/dog_rates/status/880465832...</a>	12.0
19	<a href="https://twitter.com/dog_rates/status/880221127...">https://twitter.com/dog_rates/status/880221127...</a>	12.0
20	<a href="https://twitter.com/dog_rates/status/879862464...">https://twitter.com/dog_rates/status/879862464...</a>	13.0
21	<a href="https://twitter.com/dog_rates/status/878776093...">https://twitter.com/dog_rates/status/878776093...</a>	13.0
22	<a href="https://twitter.com/dog_rates/status/878281511...">https://twitter.com/dog_rates/status/878281511...</a>	13.0
23	<a href="https://twitter.com/dog_rates/status/878057613...">https://twitter.com/dog_rates/status/878057613...</a>	14.0
24	<a href="https://twitter.com/dog_rates/status/877736472...">https://twitter.com/dog_rates/status/877736472...</a>	13.0
25	<a href="https://twitter.com/dog_rates/status/877316821...">https://twitter.com/dog_rates/status/877316821...</a>	13.0
26	<a href="https://twitter.com/dog_rates/status/877201837...">https://twitter.com/dog_rates/status/877201837...</a>	12.0

27	<a href="https://twitter.com/dog_rates/status/876838120...">https://twitter.com/dog_rates/status/876838120...</a>	12.0
28	<a href="https://twitter.com/dog_rates/status/875021211...">https://twitter.com/dog_rates/status/875021211...</a>	12.0
29	<a href="https://twitter.com/dog_rates/status/874012996...">https://twitter.com/dog_rates/status/874012996...</a>	13.0
...	...	...
1841	<a href="https://twitter.com/dog_rates/status/666437273...">https://twitter.com/dog_rates/status/666437273...</a>	7.0
1842	<a href="https://twitter.com/dog_rates/status/666435652...">https://twitter.com/dog_rates/status/666435652...</a>	10.0
1843	<a href="https://twitter.com/dog_rates/status/666430724...">https://twitter.com/dog_rates/status/666430724...</a>	6.0
1844	<a href="https://twitter.com/dog_rates/status/666428276...">https://twitter.com/dog_rates/status/666428276...</a>	7.0
1845	<a href="https://twitter.com/dog_rates/status/666421158...">https://twitter.com/dog_rates/status/666421158...</a>	12.0
1846	<a href="https://twitter.com/dog_rates/status/666418789...">https://twitter.com/dog_rates/status/666418789...</a>	10.0
1847	<a href="https://twitter.com/dog_rates/status/666407126...">https://twitter.com/dog_rates/status/666407126...</a>	7.0
1848	<a href="https://twitter.com/dog_rates/status/666396247...">https://twitter.com/dog_rates/status/666396247...</a>	9.0
1849	<a href="https://twitter.com/dog_rates/status/666373753...">https://twitter.com/dog_rates/status/666373753...</a>	11.0
1850	<a href="https://twitter.com/dog_rates/status/666362758...">https://twitter.com/dog_rates/status/666362758...</a>	6.0
1851	<a href="https://twitter.com/dog_rates/status/666353288...">https://twitter.com/dog_rates/status/666353288...</a>	8.0
1852	<a href="https://twitter.com/dog_rates/status/666345417...">https://twitter.com/dog_rates/status/666345417...</a>	10.0
1853	<a href="https://twitter.com/dog_rates/status/666337882...">https://twitter.com/dog_rates/status/666337882...</a>	9.0
1854	<a href="https://twitter.com/dog_rates/status/666273097...">https://twitter.com/dog_rates/status/666273097...</a>	11.0
1855	<a href="https://twitter.com/dog_rates/status/666268910...">https://twitter.com/dog_rates/status/666268910...</a>	10.0
1856	<a href="https://twitter.com/dog_rates/status/666102155...">https://twitter.com/dog_rates/status/666102155...</a>	11.0
1857	<a href="https://twitter.com/dog_rates/status/666099513...">https://twitter.com/dog_rates/status/666099513...</a>	8.0
1858	<a href="https://twitter.com/dog_rates/status/666094000...">https://twitter.com/dog_rates/status/666094000...</a>	9.0
1859	<a href="https://twitter.com/dog_rates/status/666082916...">https://twitter.com/dog_rates/status/666082916...</a>	6.0
1860	<a href="https://twitter.com/dog_rates/status/666073100...">https://twitter.com/dog_rates/status/666073100...</a>	10.0
1861	<a href="https://twitter.com/dog_rates/status/666071193...">https://twitter.com/dog_rates/status/666071193...</a>	9.0
1862	<a href="https://twitter.com/dog_rates/status/666063827...">https://twitter.com/dog_rates/status/666063827...</a>	10.0
1863	<a href="https://twitter.com/dog_rates/status/666058600...">https://twitter.com/dog_rates/status/666058600...</a>	8.0
1864	<a href="https://twitter.com/dog_rates/status/666057090...">https://twitter.com/dog_rates/status/666057090...</a>	9.0
1865	<a href="https://twitter.com/dog_rates/status/666055525...">https://twitter.com/dog_rates/status/666055525...</a>	10.0
1866	<a href="https://twitter.com/dog_rates/status/666050758...">https://twitter.com/dog_rates/status/666050758...</a>	10.0
1867	<a href="https://twitter.com/dog_rates/status/666044226...">https://twitter.com/dog_rates/status/666044226...</a>	6.0
1868	<a href="https://twitter.com/dog_rates/status/666033412...">https://twitter.com/dog_rates/status/666033412...</a>	9.0
1869	<a href="https://twitter.com/dog_rates/status/666029285...">https://twitter.com/dog_rates/status/666029285...</a>	7.0
1870	<a href="https://twitter.com/dog_rates/status/666020888...">https://twitter.com/dog_rates/status/666020888...</a>	8.0

	rating_denominator	name	fancy_class \
0	10	Franklin	NaN
1	10	Jax	NaN
2	10	None	NaN
3	10	Koda	NaN
4	10	Ted	NaN
5	10	Zeke	NaN
6	10	Ralphus	NaN
7	10	Gerald	NaN
8	10	Canela	NaN
9	10	Maya	NaN
10	10	Mingus	NaN
11	10	Roscoe	pupper

12	10	None	NaN
13	10	Lola	NaN
14	10	None	NaN
15	10	Rusty	NaN
16	10	Alfy	NaN
17	10	Koko	NaN
18	10	Bella	NaN
19	10	Jesse	NaN
20	10	Romeo	NaN
21	10	Snoopy	puppo
22	10	Shadow	NaN
23	10	Emmy	NaN
24	10	Aja	NaN
25	10	Dante	NaN
26	10	Nelly	NaN
27	10	Ginger	pupper
28	10	None	NaN
29	10	Sebastian	puppo
...	...	...	...
1841	10	None	NaN
1842	10	None	NaN
1843	10	None	NaN
1844	10	None	NaN
1845	10	None	NaN
1846	10	Walter	NaN
1847	10	NaN	NaN
1848	10	None	NaN
1849	10	None	NaN
1850	10	None	NaN
1851	10	None	NaN
1852	10	None	NaN
1853	10	NaN	NaN
1854	10	None	NaN
1855	10	None	NaN
1856	10	None	NaN
1857	10	None	NaN
1858	10	None	NaN
1859	10	None	NaN
1860	10	None	NaN
1861	10	None	NaN
1862	10	the	NaN
1863	10	the	NaN
1864	10	NaN	NaN
1865	10	NaN	NaN
1866	10	NaN	NaN
1867	10	NaN	NaN
1868	10	NaN	NaN
1869	10	NaN	NaN

1870

10

None

NaN

	jpg_url	pic_conf	\
0	https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg	0.555712	
1	https://pbs.twimg.com/media/DF1eOmZXUAAALUcq.jpg	0.341703	
2	https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg	0.566142	
3	https://pbs.twimg.com/media/DFnwSY4WAAAMliS.jpg	0.957979	
4	https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg	0.991650	
5	https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg	0.469760	
6	https://pbs.twimg.com/media/DFTH_0-UQAACu20.jpg	0.700377	
7	https://pbs.twimg.com/media/DFMWn56WsAAkA7B.jpg	0.995026	
8	https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg	0.809197	
9	https://pbs.twimg.com/media/DE8yicJW0AAAavBJ.jpg	0.793469	
10	https://pbs.twimg.com/media/DE5Se8FXcAAJFx4.jpg	0.309706	
11	https://pbs.twimg.com/media/DE0BTnQUwAAPKEH.jpg	0.999201	
12	https://pbs.twimg.com/media/DEi_N9qXYAAgEEw.jpg	0.812482	
13	https://pbs.twimg.com/media/DEe2tZXXkAAwyX3.jpg	0.822103	
14	https://pbs.twimg.com/media/DEF2-_hXoAAAs62q.jpg	0.949562	
15	https://pbs.twimg.com/media/DEEEEnIqXYAAiJh_.jpg	0.466778	
16	https://pbs.twimg.com/media/DD5yKdPW0AAArzX8.jpg	0.762211	
17	https://pbs.twimg.com/media/DD2oCl2WAAEI_4a.jpg	0.949591	
18	https://pbs.twimg.com/media/DDgK-J4XUAIEV9W.jpg	0.913255	
19	https://pbs.twimg.com/media/DDcscbXU0AIfDzs.jpg	0.238525	
20	https://pbs.twimg.com/media/DDXmPrbWAAEKMy.jpg	0.813507	
21	https://pbs.twimg.com/media/DDIKMXzW0AEibje.jpg	0.734684	
22	https://pbs.twimg.com/media/DDBIX9QVYAAohGa.jpg	0.320420	
23	https://pbs.twimg.com/media/DC98vABUIAA97pz.jpg	0.839097	
24	https://pbs.twimg.com/media/DC5YqoQW0AAArOLH.jpg	0.837956	
25	https://pbs.twimg.com/media/DCza_vtXkAQXGpC.jpg	0.509967	
26	https://pbs.twimg.com/media/DCxyahJW5AAAddSC.jpg	0.931120	
27	https://pbs.twimg.com/media/DCsnnZsVwAEfkyi.jpg	0.575751	
28	https://pbs.twimg.com/media/DCSzF3NV0AAPzT4.jpg	0.714319	
29	https://pbs.twimg.com/media/DCEeLxjXsAAvNSM.jpg	0.806674	
...	...	...	...
1841	https://pbs.twimg.com/media/CT-pKmRWIAAxUWj.jpg	0.671853	
1842	https://pbs.twimg.com/media/CT-nsTQWEAEkyDn.jpg	0.184130	
1843	https://pbs.twimg.com/media/CT-jNYqW4AAPi2M.jpg	0.505184	
1844	https://pbs.twimg.com/media/CT-g-ODUwAEQdSn.jpg	0.371361	
1845	https://pbs.twimg.com/media/CT-aggCXAAIMfT3.jpg	0.906777	
1846	https://pbs.twimg.com/media/CT-YWb7U8AA7QnN.jpg	0.149680	
1847	https://pbs.twimg.com/media/CT-NvwmW4AAugGZ.jpg	0.529139	
1848	https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg	0.978108	
1849	https://pbs.twimg.com/media/CT9vZEYWUAA1ZO5.jpg	0.326467	
1850	https://pbs.twimg.com/media/CT9lXGsUcAAyUFt.jpg	0.996496	
1851	https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg	0.336874	
1852	https://pbs.twimg.com/media/CT9Vn7PW0AA_ZCM.jpg	0.858744	
1853	https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg	0.416669	
1854	https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg	0.176053	

1855	<a href="https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg">https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg</a>	0.086502
1856	<a href="https://pbs.twimg.com/media/CT54YGiWUAEZnoK.jpg">https://pbs.twimg.com/media/CT54YGiWUAEZnoK.jpg</a>	0.298617
1857	<a href="https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg">https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg</a>	0.582330
1858	<a href="https://pbs.twimg.com/media/CT5w9gUW4AAxBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAxBNN.jpg</a>	0.195217
1859	<a href="https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg">https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg</a>	0.489814
1860	<a href="https://pbs.twimg.com/media/CT5d9DZXXXXLcwe.jpg">https://pbs.twimg.com/media/CT5d9DZXXXXLcwe.jpg</a>	0.260857
1861	<a href="https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg">https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg</a>	0.503672
1862	<a href="https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg">https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg</a>	0.775930
1863	<a href="https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg">https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg</a>	0.201493
1864	<a href="https://pbs.twimg.com/media/CT5PY90WwAAQGLo.jpg">https://pbs.twimg.com/media/CT5PY90WwAAQGLo.jpg</a>	0.962465
1865	<a href="https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg">https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg</a>	0.692517
1866	<a href="https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg">https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg</a>	0.651137
1867	<a href="https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg">https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg</a>	0.408143
1868	<a href="https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg">https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg</a>	0.596461
1869	<a href="https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg">https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg</a>	0.506826
1870	<a href="https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg">https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg</a>	0.465074

	breed	favorite_count	retweet_count
0	basset	41048	9774
1	Appenzeller	12041	2158
2	Pomeranian	56848	16716
3	Samoyed	31166	7624
4	French_bulldog	27672	4735
5	golden_retriever	26080	4559
6	Siberian_husky	20290	3732
7	French_bulldog	22201	3653
8	Pembroke	69871	18781
9	Chihuahua	35859	8084
10	kuvasz	12306	3443
11	French_bulldog	21524	3316
12	malamute	22367	4556
13	chow	28514	6096
14	golden_retriever	37914	6949
15	Eskimo_dog	24445	4122
16	golden_retriever	45880	12118
17	web_site	29900	5203
18	golden_retriever	29075	6546
19	Chihuahua	27640	4436
20	basset	22667	3642
21	Italian_greyhound	19763	4319
22	basset	7913	1349
23	French_bulldog	42876	7181
24	Chesapeake_Bay_retriever	71144	17300
25	Saluki	27907	5414
26	Pembroke	27755	5880
27	bloodhound	21125	3506
28	West_Highland_white_terrier	26022	4922
29	Cardigan	35501	11007

...	...	...	...
1841	Chihuahua	131	52
1842	Chesapeake_Bay_retriever	170	54
1843	llama	330	204
1844	Pembroke	171	90
1845	Blenheim_spaniel	327	118
1846	toy_terrier	129	48
1847	black-and-tan_coonhound	113	44
1848	Chihuahua	172	92
1849	soft-coated_wheaten_terrier	194	100
1850	guinea_pig	804	595
1851	malamute	229	77
1852	golden_retriever	307	146
1853	ox	204	96
1854	Italian_greyhound	184	82
1855	desktop_computer	108	37
1856	English_setter	81	16
1857	Lhasa	164	73
1858	bloodhound	169	79
1859	pug	121	47
1860	Walker_hound	335	174
1861	Gordon_setter	154	67
1862	golden_retriever	496	232
1863	miniature_poodle	115	61
1864	shopping_cart	304	146
1865	chow	448	261
1866	Bernese_mountain_dog	136	60
1867	Rhodesian_ridgeback	311	147
1868	German_shepherd	128	47
1869	redbone	132	48
1870	Welsh_springer_spaniel	2535	532

[1871 rows x 14 columns]

## Test

In [61]: twitter\_master

Out[61]:	tweet_id	timestamp	source	\
0	891327558926688256	2017-07-29 16:00:24	twitter for iphone	
1	890971913173991426	2017-07-28 16:27:12	twitter for iphone	
2	890729181411237888	2017-07-28 00:22:40	twitter for iphone	
3	890006608113172480	2017-07-26 00:31:25	twitter for iphone	
4	889638837579907072	2017-07-25 00:10:02	twitter for iphone	
5	888804989199671297	2017-07-22 16:56:37	twitter for iphone	
6	888554962724278272	2017-07-22 00:23:06	twitter for iphone	
7	888078434458587136	2017-07-20 16:49:33	twitter for iphone	
8	887473957103951883	2017-07-19 00:47:34	twitter for iphone	

9	886983233522544640	2017-07-17	16:17:36	twitter for iphone
10	886736880519319552	2017-07-16	23:58:41	twitter for iphone
11	886366144734445568	2017-07-15	23:25:31	twitter for iphone
12	885167619883638784	2017-07-12	16:03:00	twitter for iphone
13	884876753390489601	2017-07-11	20:47:12	twitter for iphone
14	883117836046086144	2017-07-07	00:17:54	twitter for iphone
15	882992080364220416	2017-07-06	15:58:11	twitter for iphone
16	882268110199369728	2017-07-04	16:01:23	twitter for iphone
17	882045870035918850	2017-07-04	01:18:17	twitter for iphone
18	880465832366813184	2017-06-29	16:39:47	twitter for iphone
19	880221127280381952	2017-06-29	00:27:25	twitter for iphone
20	879862464715927552	2017-06-28	00:42:13	twitter for iphone
21	878776093423087618	2017-06-25	00:45:22	twitter for iphone
22	878281511006478336	2017-06-23	16:00:04	twitter for iphone
23	878057613040115712	2017-06-23	01:10:23	twitter for iphone
24	877736472329191424	2017-06-22	03:54:17	twitter for iphone
25	877316821321428993	2017-06-21	00:06:44	twitter for iphone
26	877201837425926144	2017-06-20	16:29:50	twitter for iphone
27	876838120628539392	2017-06-19	16:24:33	twitter for iphone
28	875021211251597312	2017-06-14	16:04:48	twitter for iphone
29	874012996292530176	2017-06-11	21:18:31	twitter for iphone
...	...	...	...	...
1841	666437273139982337	2015-11-17	02:06:42	twitter for iphone
1842	666435652385423360	2015-11-17	02:00:15	twitter for iphone
1843	666430724426358785	2015-11-17	01:40:41	twitter for iphone
1844	666428276349472768	2015-11-17	01:30:57	twitter for iphone
1845	666421158376562688	2015-11-17	01:02:40	twitter for iphone
1846	666418789513326592	2015-11-17	00:53:15	twitter for iphone
1847	666407126856765440	2015-11-17	00:06:54	twitter for iphone
1848	666396247373291520	2015-11-16	23:23:41	twitter for iphone
1849	666373753744588802	2015-11-16	21:54:18	twitter for iphone
1850	666362758909284353	2015-11-16	21:10:36	twitter for iphone
1851	666353288456101888	2015-11-16	20:32:58	twitter for iphone
1852	666345417576210432	2015-11-16	20:01:42	twitter for iphone
1853	666337882303524864	2015-11-16	19:31:45	twitter for iphone
1854	666273097616637952	2015-11-16	15:14:19	twitter for iphone
1855	666268910803644416	2015-11-16	14:57:41	twitter for iphone
1856	666102155909144576	2015-11-16	03:55:04	twitter for iphone
1857	666099513787052032	2015-11-16	03:44:34	twitter for iphone
1858	666094000022159362	2015-11-16	03:22:39	twitter for iphone
1859	666082916733198337	2015-11-16	02:38:37	twitter for iphone
1860	666073100786774016	2015-11-16	01:59:36	twitter for iphone
1861	666071193221509120	2015-11-16	01:52:02	twitter for iphone
1862	666063827256086533	2015-11-16	01:22:45	twitter for iphone
1863	666058600524156928	2015-11-16	01:01:59	twitter for iphone
1864	666057090499244032	2015-11-16	00:55:59	twitter for iphone
1865	666055525042405380	2015-11-16	00:49:46	twitter for iphone
1866	666050758794694657	2015-11-16	00:30:50	twitter for iphone



1867 666044226329800704 2015-11-16 00:04:52 twitter for iphone  
 1868 666033412701032449 2015-11-15 23:21:54 twitter for iphone  
 1869 666029285002620928 2015-11-15 23:05:30 twitter for iphone  
 1870 666020888022790149 2015-11-15 22:32:08 twitter for iphone

text \

0 This is Franklin. He would like you to stop ca...  
 1 Meet Jax. He enjoys ice cream so much he gets ...  
 2 When you watch your owner call another dog a g...  
 3 This is Koda. He is a South Australian decksha...  
 4 This is Ted. He does his best. Sometimes that'...  
 5 This is Zeke. He has a new stick. Very proud o...  
 6 This is Ralphus. He's powering up. Attempting ...  
 7 This is Gerald. He was just told he didn't get...  
 8 This is Canela. She attempted some fancy porch...  
 9 This is Maya. She's very shy. Rarely leaves he...  
 10 This is Mingus. He's a wonderful father to his...  
 11 This is Roscoe. Another pupper fallen victim t...  
 12 Here we have a corgi undercover as a malamute...  
 13 This is Lola. It's her first time outside. Mus...  
 14 Please only send dogs. We don't rate mechanics...  
 15 This is Rusty. He wasn't ready for the first p...  
 16 This is Alf. You're witnessing his first wate...  
 17 This is Koko. Her owner, inspired by Barney, r...  
 18 This is Bella. She had her first beach experie...  
 19 Meet Jesse. He's a Fetty Woof. His tongue ejec...  
 20 This is Romeo. He would like to do an entrance...  
 21 This is Snoopy. He's a proud #PrideMonthPuppo...  
 22 Meet Shadow. In an attempt to reach maximum zo...  
 23 This is Emmy. She was adopted today. Massive r...  
 24 This is Aja. She was just told she's a good do...  
 25 Meet Dante. At first he wasn't a fan of his ne...  
 26 This is Nelly. He graduated with his dogtorate...  
 27 This is Ginger. She's having a ruff Monday. To...  
 28 Guys please stop sending pictures without any ...  
 29 This is Sebastian. He can't see all the colors...  
 ...  
 1841 Here we see a lone northeastern Cumberbatch. H...  
 1842 "Can you behave? You're ruining my wedding day...  
 1843 Oh boy what a pup! Sunglasses take this one to...  
 1844 Here we have an Austrian Pulitzer. Collectors ...  
 1845 \*internally screaming\* 12/10  
 1846 This is Walter. He is an Alaskan Terrapin. Lov...  
 1847 This is a southern Vesuvius bumblegruff. Can d...  
 1848 Oh goodness. A super rare northeast Qdoba kang...  
 1849 Those are sunglasses and a jean jacket. 11/10 ...  
 1850 Unique dog here. Very small. Lives in containe...  
 1851 Here we have a mixed Asiago from the Galápagos...

1852 Look at this jokester thinking seat belt laws ...  
 1853 This is an extremely rare horned Parthenon. No...  
 1854 Can take selfies 11/10  
 1855 Very concerned about fellow dog trapped in com...  
 1856 Oh my. Here you are seeing an Adobe Setter giv...  
 1857 Can stand on stump for what seems like a while...  
 1858 This appears to be a Mongolian Presbyterian mi...  
 1859 Here we have a well-established sunblockerspan...  
 1860 Let's hope this flight isn't Malaysian (lol). ...  
 1861 Here we have a northern speckled Rhododendron...  
 1862 This is the happiest dog you will ever see. Ve...  
 1863 Here is the Rand Paul of retrievers folks! He'...  
 1864 My oh my. This is a rare blond Canadian terrie...  
 1865 Here is a Siberian heavily armored polar bear ...  
 1866 This is a truly beautiful English Wilson Staff...  
 1867 This is a purebred Piers Morgan. Loves to Netf...  
 1868 Here is a very happy pup. Big fan of well-main...  
 1869 This is a western brown Mitsubishi terrier. Up...  
 1870 Here we have a Japanese Irish Setter. Lost eye...

	expanded_urls	rating_numerator \
0	<a href="https://twitter.com/dog_rates/status/891327558...">https://twitter.com/dog_rates/status/891327558...</a>	12.0
1	<a href="https://twitter.com/dog_rates/status/890971913...">https://twitter.com/dog_rates/status/890971913...</a>	13.0
2	<a href="https://twitter.com/dog_rates/status/890729181...">https://twitter.com/dog_rates/status/890729181...</a>	13.0
3	<a href="https://twitter.com/dog_rates/status/890006608...">https://twitter.com/dog_rates/status/890006608...</a>	13.0
4	<a href="https://twitter.com/dog_rates/status/889638837...">https://twitter.com/dog_rates/status/889638837...</a>	12.0
5	<a href="https://twitter.com/dog_rates/status/888804989...">https://twitter.com/dog_rates/status/888804989...</a>	13.0
6	<a href="https://twitter.com/dog_rates/status/888554962...">https://twitter.com/dog_rates/status/888554962...</a>	13.0
7	<a href="https://twitter.com/dog_rates/status/888078434...">https://twitter.com/dog_rates/status/888078434...</a>	12.0
8	<a href="https://twitter.com/dog_rates/status/887473957...">https://twitter.com/dog_rates/status/887473957...</a>	13.0
9	<a href="https://twitter.com/dog_rates/status/886983233...">https://twitter.com/dog_rates/status/886983233...</a>	13.0
10	<a href="https://twitter.com/dog_rates/status/886736880...">https://twitter.com/dog_rates/status/886736880...</a>	13.0
11	<a href="https://twitter.com/dog_rates/status/886366144...">https://twitter.com/dog_rates/status/886366144...</a>	12.0
12	<a href="https://twitter.com/dog_rates/status/885167619...">https://twitter.com/dog_rates/status/885167619...</a>	13.0
13	<a href="https://twitter.com/dog_rates/status/884876753...">https://twitter.com/dog_rates/status/884876753...</a>	13.0
14	<a href="https://twitter.com/dog_rates/status/883117836...">https://twitter.com/dog_rates/status/883117836...</a>	13.0
15	<a href="https://twitter.com/dog_rates/status/882992080...">https://twitter.com/dog_rates/status/882992080...</a>	13.0
16	<a href="https://twitter.com/dog_rates/status/882268110...">https://twitter.com/dog_rates/status/882268110...</a>	13.0
17	<a href="https://twitter.com/dog_rates/status/882045870...">https://twitter.com/dog_rates/status/882045870...</a>	13.0
18	<a href="https://twitter.com/dog_rates/status/880465832...">https://twitter.com/dog_rates/status/880465832...</a>	12.0
19	<a href="https://twitter.com/dog_rates/status/880221127...">https://twitter.com/dog_rates/status/880221127...</a>	12.0
20	<a href="https://twitter.com/dog_rates/status/879862464...">https://twitter.com/dog_rates/status/879862464...</a>	13.0
21	<a href="https://twitter.com/dog_rates/status/878776093...">https://twitter.com/dog_rates/status/878776093...</a>	13.0
22	<a href="https://twitter.com/dog_rates/status/878281511...">https://twitter.com/dog_rates/status/878281511...</a>	13.0
23	<a href="https://twitter.com/dog_rates/status/878057613...">https://twitter.com/dog_rates/status/878057613...</a>	14.0
24	<a href="https://twitter.com/dog_rates/status/877736472...">https://twitter.com/dog_rates/status/877736472...</a>	13.0
25	<a href="https://twitter.com/dog_rates/status/877316821...">https://twitter.com/dog_rates/status/877316821...</a>	13.0
26	<a href="https://twitter.com/dog_rates/status/877201837...">https://twitter.com/dog_rates/status/877201837...</a>	12.0

27	<a href="https://twitter.com/dog_rates/status/876838120...">https://twitter.com/dog_rates/status/876838120...</a>	12.0
28	<a href="https://twitter.com/dog_rates/status/875021211...">https://twitter.com/dog_rates/status/875021211...</a>	12.0
29	<a href="https://twitter.com/dog_rates/status/874012996...">https://twitter.com/dog_rates/status/874012996...</a>	13.0
...	...	...
1841	<a href="https://twitter.com/dog_rates/status/666437273...">https://twitter.com/dog_rates/status/666437273...</a>	7.0
1842	<a href="https://twitter.com/dog_rates/status/666435652...">https://twitter.com/dog_rates/status/666435652...</a>	10.0
1843	<a href="https://twitter.com/dog_rates/status/666430724...">https://twitter.com/dog_rates/status/666430724...</a>	6.0
1844	<a href="https://twitter.com/dog_rates/status/666428276...">https://twitter.com/dog_rates/status/666428276...</a>	7.0
1845	<a href="https://twitter.com/dog_rates/status/666421158...">https://twitter.com/dog_rates/status/666421158...</a>	12.0
1846	<a href="https://twitter.com/dog_rates/status/666418789...">https://twitter.com/dog_rates/status/666418789...</a>	10.0
1847	<a href="https://twitter.com/dog_rates/status/666407126...">https://twitter.com/dog_rates/status/666407126...</a>	7.0
1848	<a href="https://twitter.com/dog_rates/status/666396247...">https://twitter.com/dog_rates/status/666396247...</a>	9.0
1849	<a href="https://twitter.com/dog_rates/status/666373753...">https://twitter.com/dog_rates/status/666373753...</a>	11.0
1850	<a href="https://twitter.com/dog_rates/status/666362758...">https://twitter.com/dog_rates/status/666362758...</a>	6.0
1851	<a href="https://twitter.com/dog_rates/status/666353288...">https://twitter.com/dog_rates/status/666353288...</a>	8.0
1852	<a href="https://twitter.com/dog_rates/status/666345417...">https://twitter.com/dog_rates/status/666345417...</a>	10.0
1853	<a href="https://twitter.com/dog_rates/status/666337882...">https://twitter.com/dog_rates/status/666337882...</a>	9.0
1854	<a href="https://twitter.com/dog_rates/status/666273097...">https://twitter.com/dog_rates/status/666273097...</a>	11.0
1855	<a href="https://twitter.com/dog_rates/status/666268910...">https://twitter.com/dog_rates/status/666268910...</a>	10.0
1856	<a href="https://twitter.com/dog_rates/status/666102155...">https://twitter.com/dog_rates/status/666102155...</a>	11.0
1857	<a href="https://twitter.com/dog_rates/status/666099513...">https://twitter.com/dog_rates/status/666099513...</a>	8.0
1858	<a href="https://twitter.com/dog_rates/status/666094000...">https://twitter.com/dog_rates/status/666094000...</a>	9.0
1859	<a href="https://twitter.com/dog_rates/status/666082916...">https://twitter.com/dog_rates/status/666082916...</a>	6.0
1860	<a href="https://twitter.com/dog_rates/status/666073100...">https://twitter.com/dog_rates/status/666073100...</a>	10.0
1861	<a href="https://twitter.com/dog_rates/status/666071193...">https://twitter.com/dog_rates/status/666071193...</a>	9.0
1862	<a href="https://twitter.com/dog_rates/status/666063827...">https://twitter.com/dog_rates/status/666063827...</a>	10.0
1863	<a href="https://twitter.com/dog_rates/status/666058600...">https://twitter.com/dog_rates/status/666058600...</a>	8.0
1864	<a href="https://twitter.com/dog_rates/status/666057090...">https://twitter.com/dog_rates/status/666057090...</a>	9.0
1865	<a href="https://twitter.com/dog_rates/status/666055525...">https://twitter.com/dog_rates/status/666055525...</a>	10.0
1866	<a href="https://twitter.com/dog_rates/status/666050758...">https://twitter.com/dog_rates/status/666050758...</a>	10.0
1867	<a href="https://twitter.com/dog_rates/status/666044226...">https://twitter.com/dog_rates/status/666044226...</a>	6.0
1868	<a href="https://twitter.com/dog_rates/status/666033412...">https://twitter.com/dog_rates/status/666033412...</a>	9.0
1869	<a href="https://twitter.com/dog_rates/status/666029285...">https://twitter.com/dog_rates/status/666029285...</a>	7.0
1870	<a href="https://twitter.com/dog_rates/status/666020888...">https://twitter.com/dog_rates/status/666020888...</a>	8.0

	rating_denominator	name	fancy_class \
0	10	Franklin	NaN
1	10	Jax	NaN
2	10	None	NaN
3	10	Koda	NaN
4	10	Ted	NaN
5	10	Zeke	NaN
6	10	Ralphus	NaN
7	10	Gerald	NaN
8	10	Canela	NaN
9	10	Maya	NaN
10	10	Mingus	NaN
11	10	Roscoe	pupper

12	10	None	NaN
13	10	Lola	NaN
14	10	None	NaN
15	10	Rusty	NaN
16	10	Alfy	NaN
17	10	Koko	NaN
18	10	Bella	NaN
19	10	Jesse	NaN
20	10	Romeo	NaN
21	10	Snoopy	puppo
22	10	Shadow	NaN
23	10	Emmy	NaN
24	10	Aja	NaN
25	10	Dante	NaN
26	10	Nelly	NaN
27	10	Ginger	pupper
28	10	None	NaN
29	10	Sebastian	puppo
...	...	...	...
1841	10	None	NaN
1842	10	None	NaN
1843	10	None	NaN
1844	10	None	NaN
1845	10	None	NaN
1846	10	Walter	NaN
1847	10	NaN	NaN
1848	10	None	NaN
1849	10	None	NaN
1850	10	None	NaN
1851	10	None	NaN
1852	10	None	NaN
1853	10	NaN	NaN
1854	10	None	NaN
1855	10	None	NaN
1856	10	None	NaN
1857	10	None	NaN
1858	10	None	NaN
1859	10	None	NaN
1860	10	None	NaN
1861	10	None	NaN
1862	10	the	NaN
1863	10	the	NaN
1864	10	NaN	NaN
1865	10	NaN	NaN
1866	10	NaN	NaN
1867	10	NaN	NaN
1868	10	NaN	NaN
1869	10	NaN	NaN

1870

10

None

NaN

	jpg_url	pic_conf	\
0	https://pbs.twimg.com/media/DF6hr6BUMAAzZgT.jpg	0.555712	
1	https://pbs.twimg.com/media/DF1eOmZXUAAALUcq.jpg	0.341703	
2	https://pbs.twimg.com/media/DFyBahAVwAAhUTd.jpg	0.566142	
3	https://pbs.twimg.com/media/DFnwSY4WAAAMliS.jpg	0.957979	
4	https://pbs.twimg.com/media/DFihzFfXsAYGDPR.jpg	0.991650	
5	https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg	0.469760	
6	https://pbs.twimg.com/media/DFTH_0-UQAACu20.jpg	0.700377	
7	https://pbs.twimg.com/media/DFMWn56WsAAkA7B.jpg	0.995026	
8	https://pbs.twimg.com/media/DFDw2tyUQAAAFke.jpg	0.809197	
9	https://pbs.twimg.com/media/DE8yicJW0AAAavBJ.jpg	0.793469	
10	https://pbs.twimg.com/media/DE5Se8FXcAAJFx4.jpg	0.309706	
11	https://pbs.twimg.com/media/DE0BTnQUwAApKEH.jpg	0.999201	
12	https://pbs.twimg.com/media/DEi_N9qXYAAgEEw.jpg	0.812482	
13	https://pbs.twimg.com/media/DEe2tZXXkAAwyX3.jpg	0.822103	
14	https://pbs.twimg.com/media/DEF2-_hXoAAAs62q.jpg	0.949562	
15	https://pbs.twimg.com/media/DEEEenIqXYAAiJh_.jpg	0.466778	
16	https://pbs.twimg.com/media/DD5yKdPW0AAArzX8.jpg	0.762211	
17	https://pbs.twimg.com/media/DD2oCl2WAAEI_4a.jpg	0.949591	
18	https://pbs.twimg.com/media/DDgK-J4XUAIEV9W.jpg	0.913255	
19	https://pbs.twimg.com/media/DDcscbXU0AIfDzs.jpg	0.238525	
20	https://pbs.twimg.com/media/DDXmPrbWAAEKmvy.jpg	0.813507	
21	https://pbs.twimg.com/media/DDIKMXzW0AEibje.jpg	0.734684	
22	https://pbs.twimg.com/media/DDBIX9QVYAAohGa.jpg	0.320420	
23	https://pbs.twimg.com/media/DC98vABUIAA97pz.jpg	0.839097	
24	https://pbs.twimg.com/media/DC5YqoQW0AAArOLH.jpg	0.837956	
25	https://pbs.twimg.com/media/DCza_vtXkAQXGpC.jpg	0.509967	
26	https://pbs.twimg.com/media/DCxyahJWsAAAddSC.jpg	0.931120	
27	https://pbs.twimg.com/media/DCsnnZsVwAEfkyi.jpg	0.575751	
28	https://pbs.twimg.com/media/DCSzF3NV0AAPzT4.jpg	0.714319	
29	https://pbs.twimg.com/media/DCEeLxjXsAAvNSM.jpg	0.806674	
...	...	...	...
1841	https://pbs.twimg.com/media/CT-pKmRWIAAxUWj.jpg	0.671853	
1842	https://pbs.twimg.com/media/CT-nsTQWEAEkyDn.jpg	0.184130	
1843	https://pbs.twimg.com/media/CT-jNYqW4AAPi2M.jpg	0.505184	
1844	https://pbs.twimg.com/media/CT-g-ODUwAEQdSn.jpg	0.371361	
1845	https://pbs.twimg.com/media/CT-aggCXAAIMfT3.jpg	0.906777	
1846	https://pbs.twimg.com/media/CT-YWb7U8AA7QnN.jpg	0.149680	
1847	https://pbs.twimg.com/media/CT-NvwmW4AAAugGZ.jpg	0.529139	
1848	https://pbs.twimg.com/media/CT-D2ZHWIAA3gK1.jpg	0.978108	
1849	https://pbs.twimg.com/media/CT9vZEYWUAA1ZO5.jpg	0.326467	
1850	https://pbs.twimg.com/media/CT9lXGsUcAAyUFt.jpg	0.996496	
1851	https://pbs.twimg.com/media/CT9cx0tUEAAhNN_.jpg	0.336874	
1852	https://pbs.twimg.com/media/CT9Vn7PW0AA_ZCM.jpg	0.858744	
1853	https://pbs.twimg.com/media/CT90wFIWEAMuRje.jpg	0.416669	
1854	https://pbs.twimg.com/media/CT8T1mtUwAA3aqm.jpg	0.176053	

1855	<a href="https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg">https://pbs.twimg.com/media/CT8QCd1WEAADXws.jpg</a>	0.086502
1856	<a href="https://pbs.twimg.com/media/CT54YGiWUAEZnoK.jpg">https://pbs.twimg.com/media/CT54YGiWUAEZnoK.jpg</a>	0.298617
1857	<a href="https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg">https://pbs.twimg.com/media/CT51-JJUEAA6hV8.jpg</a>	0.582330
1858	<a href="https://pbs.twimg.com/media/CT5w9gUW4AAxBNN.jpg">https://pbs.twimg.com/media/CT5w9gUW4AAxBNN.jpg</a>	0.195217
1859	<a href="https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg">https://pbs.twimg.com/media/CT5m4VGWEAAAtKc8.jpg</a>	0.489814
1860	<a href="https://pbs.twimg.com/media/CT5d9DZXAALcwe.jpg">https://pbs.twimg.com/media/CT5d9DZXAALcwe.jpg</a>	0.260857
1861	<a href="https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg">https://pbs.twimg.com/media/CT5cN_3WEAA10oZ.jpg</a>	0.503672
1862	<a href="https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg">https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg</a>	0.775930
1863	<a href="https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg">https://pbs.twimg.com/media/CT5Qw94XAAA_2dP.jpg</a>	0.201493
1864	<a href="https://pbs.twimg.com/media/CT5PY90WwAAQGLo.jpg">https://pbs.twimg.com/media/CT5PY90WwAAQGLo.jpg</a>	0.962465
1865	<a href="https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg">https://pbs.twimg.com/media/CT5N9tpXIAAifs1.jpg</a>	0.692517
1866	<a href="https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg">https://pbs.twimg.com/media/CT5Jof1WUAEuVxN.jpg</a>	0.651137
1867	<a href="https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg">https://pbs.twimg.com/media/CT5Dr8HUEAA-lEu.jpg</a>	0.408143
1868	<a href="https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg">https://pbs.twimg.com/media/CT4521TWwAEvMyu.jpg</a>	0.596461
1869	<a href="https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg">https://pbs.twimg.com/media/CT42GRgUYAA5iDo.jpg</a>	0.506826
1870	<a href="https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg">https://pbs.twimg.com/media/CT4udnOWwAA0aMy.jpg</a>	0.465074

	breed	favorite_count	retweet_count
0	basset	41048	9774
1	Appenzeller	12041	2158
2	Pomeranian	56848	16716
3	Samoyed	31166	7624
4	French_bulldog	27672	4735
5	golden_retriever	26080	4559
6	Siberian_husky	20290	3732
7	French_bulldog	22201	3653
8	Pembroke	69871	18781
9	Chihuahua	35859	8084
10	kuvasz	12306	3443
11	French_bulldog	21524	3316
12	malamute	22367	4556
13	chow	28514	6096
14	golden_retriever	37914	6949
15	Eskimo_dog	24445	4122
16	golden_retriever	45880	12118
17	web_site	29900	5203
18	golden_retriever	29075	6546
19	Chihuahua	27640	4436
20	basset	22667	3642
21	Italian_greyhound	19763	4319
22	basset	7913	1349
23	French_bulldog	42876	7181
24	Chesapeake_Bay_retriever	71144	17300
25	Saluki	27907	5414
26	Pembroke	27755	5880
27	bloodhound	21125	3506
28	West_Highland_white_terrier	26022	4922
29	Cardigan	35501	11007

...	...	...	...
1841	Chihuahua	131	52
1842	Chesapeake_Bay_retriever	170	54
1843	llama	330	204
1844	Pembroke	171	90
1845	Blenheim_spaniel	327	118
1846	toy_terrier	129	48
1847	black-and-tan_coonhound	113	44
1848	Chihuahua	172	92
1849	soft-coated_wheaten_terrier	194	100
1850	guinea_pig	804	595
1851	malamute	229	77
1852	golden_retriever	307	146
1853	ox	204	96
1854	Italian_greyhound	184	82
1855	desktop_computer	108	37
1856	English_setter	81	16
1857	Lhasa	164	73
1858	bloodhound	169	79
1859	pug	121	47
1860	Walker_hound	335	174
1861	Gordon_setter	154	67
1862	golden_retriever	496	232
1863	miniature_poodle	115	61
1864	shopping_cart	304	146
1865	chow	448	261
1866	Bernese_mountain_dog	136	60
1867	Rhodesian_ridgeback	311	147
1868	German_shepherd	128	47
1869	redbone	132	48
1870	Welsh_springer_spaniel	2535	532

[1871 rows x 14 columns]

```
In [63]: twitter_master.to_csv('twitter_master.csv')
```

```
In [ ]:
```