

Untitled

December 12, 2018

1 Analyzing, and Visualizing Data

We Rate Dogs dataset after wrangling.. By Philipp Keupp

```
In [1]: import matplotlib
import matplotlib.pyplot as plt
import pandas as pd
import datetime as dt
import seaborn as sns
```

```
%matplotlib inline
```

```
In [2]: # Change the style of the plots (http://tonysyu.github.io/raw\_content/matplotlib-style)
matplotlib.style.use('ggplot')
```

```
In [3]: # Import the clean dataset into dataframe
df_master = pd.read_csv('twitter_archive_master.csv')
df_master.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 1625 entries, 0 to 1624
```

```
Data columns (total 15 columns):
```

tweet_id	1625 non-null int64
text	1625 non-null object
rating_numerator	1625 non-null float64
rating_denominator	1625 non-null float64
name	1625 non-null object
dog_stage	1625 non-null object
year	1625 non-null int64
month	1625 non-null int64
day	1625 non-null int64
rating	1625 non-null float64
jpg_url	1625 non-null object
dog_type	1625 non-null object
confidence_list	1625 non-null float64
favorite_count	1624 non-null float64
retweet_count	1624 non-null float64
dtypes: float64(6), int64(4), object(5)	
memory usage: 190.5+ KB	

1.1 Introduction

First, WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators however are almost always greater than 10 (11/10, 12/10, 13/10, etc.). In theory the rates should be 1 to 10. However, WeRateDogs disagree, they admit almost all dogs deserve a 10 and sometimes more than that.

WeRateDogs has over 6 million followers and has received international media coverage. One of those coverage were about the quote "they're good dogs Brent", this was an exchange in which WeRateDogs shut down a person having an issue with its rating system in humorous ways.

So what does this rating do? Which dog type (breed) is most common dog in the tweet dataset? What dog type has the highest average rating? These questions and more are answered in the following insights:

1.2 Insight one & visualization

Golden retriever is the most common dog in this dataset.

```
In [7]: df_master['dog_type'].value_counts()
```

```
Out[7]: golden_retriever      142
         Labrador_retriever    103
         Pembroke              94
         Chihuahua             87
         pug                   60
         toy_poodle            50
         chow                  45
         Pomeranian            42
         Samoyed               38
         malamute              32
         Chesapeake_Bay_retriever 31
         French_bulldog        30
         cocker_spaniel         26
         miniature_pinscher     24
         Eskimo_dog             22
         German_shepherd        21
         Staffordshire_bullterrier 20
         Siberian_husky         20
         Shih-Tzu              20
         Rottweiler            19
         Cardigan              19
         Shetland_sheepdog      19
         Maltese_dog           19
         beagle                19
         kuvasz                18
         Italian_greyhound      17
         basset                17
         Lakeland_terrier       17
         American_Staffordshire_terrier 16
```

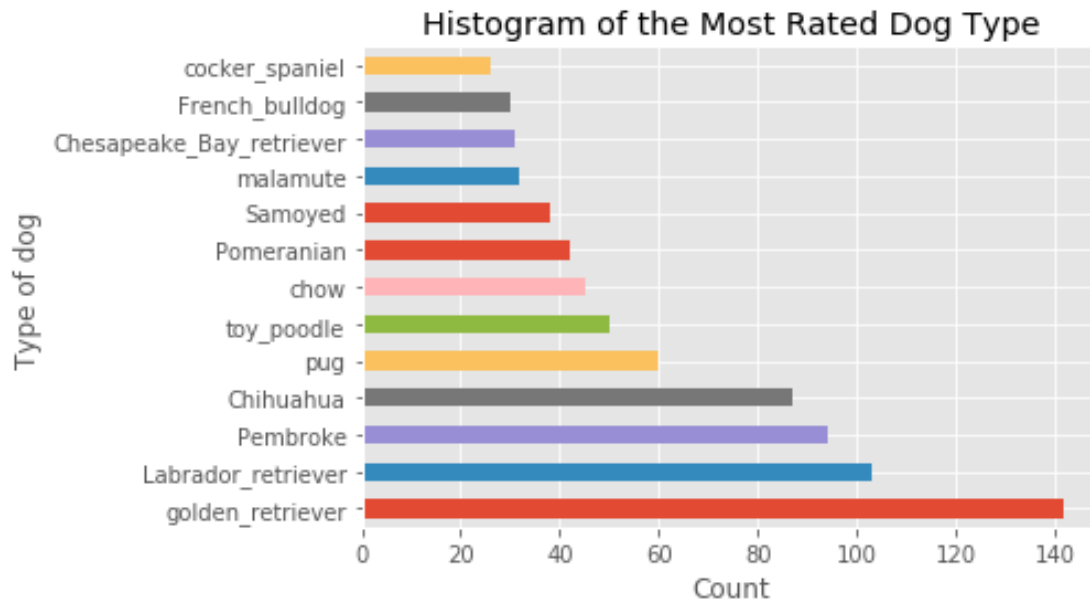
West_Highland_white_terrier	16
...	
Tibetan_terrier	4
Weimaraner	4
giant_schnauzer	4
Rhodesian_ridgeback	4
Brabancon_griffon	3
briard	3
komondor	3
Greater_Swiss_Mountain_dog	3
Irish_water_spaniel	3
toy_terrier	3
curly-coated_retriever	3
Leonberg	3
Irish_setter	3
cairn	3
Tibetan_mastiff	3
groenendael	2
Appenzeller	2
Sussex_spaniel	2
wire-haired_fox_terrier	2
black-and-tan_coonhound	2
Afghan_hound	2
Australian_terrier	2
standard_schnauzer	1
Irish_wolfhound	1
EntleBucher	1
silky_terrier	1
clumber	1
Bouvier_des_Flandres	1
Japanese_spaniel	1
Scotch_terrier	1

Name: dog_type, Length: 113, dtype: int64

```
In [9]: df_dog_type = df_master.groupby('dog_type').filter(lambda x: len(x) >= 25)
```

```
df_dog_type['dog_type'].value_counts().plot(kind = 'barh')
plt.title('Histogram of the Most Rated Dog Type')
plt.xlabel('Count')
plt.ylabel('Type of dog')

fig = plt.gcf()
fig.savefig('output.png',bbox_inches='tight');
```



WeRateDogs has over 6000+ tweets. I was able to analyzed around 1500+ tweets. The most rated dog was golden retriever with more than 140 ratings.

Here are Tweet examples:

```
In [19]: df_master[df_master['dog_type'] == 'golden_retriever']
```

```
Out[19]:
```

	tweet_id	text \
8	667530908589760512	Meet Naphaniel. He doesn't necessarily enjoy h...
30	667044094246576128	12/10 gimme now https://t.co/QZAnwgnOMB
37	667160273090932737	This is Bradley. He is a Ronaldinho Matsuyama ...
43	667182792070062081	This is Timison. He just told an awful joke bu...
46	667211855547486208	This is Genevieve. She is a golden retriever c...
62	668286279830867968	Meet Rusty. Rusty's dreaming of a world where ...
76	666345417576210432	Look at this jokester thinking seat belt laws ...
78	666063827256086533	This is the happiest dog you will ever see. Ve...
94	666057090499244032	My oh my. This is a rare blond Canadian terrie...
110	876484053909872640	This is Benedict. He wants to thank you for th...
128	888804989199671297	This is Zeke. He has a new stick. Very proud o...
141	888917238123831296	This is Jim. He found a fren. Taught him how t...
150	880465832366813184	This is Bella. She had her first beach experie...
154	882045870035918850	This is Koko. Her owner, inspired by Barney, r...
155	882268110199369728	This is Alfy. You're witnessing his first wate...
158	883117836046086144	Please only send dogs. We don't rate mechanics...
159	883482846933004288	This is Bella. She hopes her smile made you sm...
170	690735892932222976	Say hello to Peaches. She's a Dingleberry Zand...
171	690959652130045952	This golden is happy to refute the soft mouth ...
185	692568918515392513	This is Chester. He's been guarding this pumpk...
257	684225744407494656	Two sneaky puppies were not initially seen, mo...

262 683098815881154561 aahhhhhkslaldhwnxmzbbbs 12/10 for being da smoos...
 294 688385280030670848 This is Louis. He's takes top-notch selfies. 1...
 332 713900603437621249 Happy Saturday here's 9 puppers on a bench. 99...
 344 708810915978854401 This is Mister. He's a wonderful father to his...
 354 709556954897764353 This is Cecil. She's a Gigglefloof Poofers. Out...
 370 718971898235854848 This is Sadie. She is prepared for battle. 10/...
 372 719339463458033665 Say hello to Lucy and Sophie. They think they'...
 385 723673163800948736 This is Ivar. She is a badass Viking warrior. ...
 398 715680795826982913 This is Zeus. He's downright fabulous. 12/10 h...

 1376 846514051647705089 This is Barney. He's an elder doggo. Hitches a...
 1381 834574053763584002 Here's a doggo fully pupared for a shower. H*c...
 1382 719991154352222208 This doggo was initially thrilled when she saw...
 1391 716080869887381504 Here's a super majestic doggo and a sunset 11/...
 1393 795464331001561088 Elder doggo does a splash. Both 13/10 incredib...
 1395 786363235746385920 This is Rizzo. He has many talents. A true ren...
 1403 782747134529531904 This is Deacon. He's the happiest almost dry d...
 1418 858843525470990336 I have stumbled puppon a doggo painting party...
 1419 821765923262631936 This is Duchess. She uses dark doggo forces to...
 1429 800388270626521089 This is Doc. He takes time out of every day to...
 1443 724405726123311104 This is Ashleigh. She's having Coachella withd...
 1479 678341075375947776 This pupper likes tape. 12/10 <https://t.co/cSp...>
 1481 676470639084101634 Three generations of pupper. 11/10 for all htt...
 1498 680100725817409536 This pupper is not coming inside until she cat...
 1522 808106460588765185 Here we have Burke (pupper) and Dexter (doggo)...
 1524 807621403335917568 This is Ollie Vue. He was a 3 legged pupper on...
 1535 751456908746354688 Here's a pupper that's very hungry but too laz...
 1537 699801817392291840 This is Cooper. He began to tear up when his b...
 1540 699434518667751424 I know this is a tad late but here's a wonderf...
 1541 673295268553605120 Meet Eve. She's a raging alcoholic 8/10 (would...
 1551 700143752053182464 When it's Janet from accounting's birthday but...
 1553 672622327801233409 This lil pupper is sad because we haven't foun...
 1573 756939218950160384 This is Jax. He is a majestic mountain pupper...
 1579 837820167694528512 Here's a pupper before and after being asked "...
 1584 765371061932261376 This is Louie. He's had a long day. Did a lot ...
 1585 693262851218264065 I hope you guys enjoy this beautiful snowy pup...
 1588 675740360753160193 Here's a pupper licking in slow motion. 12/10 ...
 1598 759793422261743616 Meet Maggie & Lila. Maggie is the doggo, L...
 1615 859607811541651456 Sorry for the lack of posts today. I came home...
 1617 889531135344209921 This is Stuart. He's sporting his favorite fan...

	rating_numerator	rating_denominator	name	dog_stage	year	month	\
8	10.0	10.0	Naphaniel	None	2015	11	
30	12.0	10.0	None	None	2015	11	
37	11.0	10.0	Bradlay	None	2015	11	
43	10.0	10.0	Timison	None	2015	11	
46	9.0	10.0	Genevieve	None	2015	11	
62	11.0	10.0	Rusty	None	2015	11	

76	10.0	10.0	None	None	2015	11
78	10.0	10.0	the	None	2015	11
94	9.0	10.0	a	None	2015	11
110	13.0	10.0	Benedict	None	2017	6
128	13.0	10.0	Zeke	None	2017	7
141	12.0	10.0	Jim	None	2017	7
150	12.0	10.0	Bella	None	2017	6
154	13.0	10.0	Koko	None	2017	7
155	13.0	10.0	Alfy	None	2017	7
158	13.0	10.0	None	None	2017	7
159	13.5	10.0	Bella	None	2017	7
170	13.0	10.0	Peaches	None	2016	1
171	11.0	10.0	None	None	2016	1
185	12.0	10.0	Chester	None	2016	1
257	143.0	130.0	None	None	2016	1
262	12.0	10.0	None	None	2016	1
294	12.0	10.0	Louis	None	2016	1
332	99.0	90.0	None	None	2016	3
344	10.0	10.0	Mister	None	2016	3
354	12.0	10.0	Cecil	None	2016	3
370	10.0	10.0	Sadie	None	2016	4
372	10.0	10.0	Lucy	None	2016	4
385	10.0	10.0	Ivar	None	2016	4
398	12.0	10.0	Zeus	None	2016	3
...
1376	13.0	10.0	Barney	doggo	2017	3
1381	13.0	10.0	None	doggo	2017	2
1382	10.0	10.0	None	doggo	2016	4
1391	11.0	10.0	None	doggo	2016	4
1393	13.0	10.0	None	doggo	2016	11
1395	13.0	10.0	Rizzo	doggo	2016	10
1403	11.0	10.0	Deacon	doggo	2016	10
1418	13.0	10.0	None	doggo	2017	5
1419	13.0	10.0	Duchess	doggo	2017	1
1429	12.0	10.0	Doc	floofer	2016	11
1443	10.0	10.0	Ashleigh	pupper	2016	4
1479	12.0	10.0	None	pupper	2015	12
1481	11.0	10.0	None	pupper	2015	12
1498	11.0	10.0	None	pupper	2015	12
1522	12.0	10.0	None	pupper	2016	12
1524	14.0	10.0	Ollie	pupper	2016	12
1535	12.0	10.0	None	pupper	2016	7
1537	11.0	10.0	Cooper	pupper	2016	2
1540	12.0	10.0	None	pupper	2016	2
1541	8.0	10.0	Eve	pupper	2015	12
1551	10.0	10.0	None	pupper	2016	2
1553	12.0	10.0	None	pupper	2015	12
1573	12.0	10.0	Jax	pupper	2016	7

1579	12.0	10.0	None	pupper	2017	3
1584	11.0	10.0	Louie	pupper	2016	8
1585	11.0	10.0	None	pupper	2016	1
1588	12.0	10.0	None	pupper	2015	12
1598	12.0	10.0	Maggie	pupper	2016	7
1615	13.0	10.0	None	puppo	2017	5
1617	13.0	10.0	Stuart	puppo	2017	7

	day	rating	jpg_url \
8	20	10.0	https://pbs.twimg.com/media/CUOL0uGUkAAx7yh.jpg
30	18	12.0	https://pbs.twimg.com/media/CUHREBXXAAE6A9b.jpg
37	19	11.0	https://pbs.twimg.com/media/CUI6uuaW4AAvCIs.jpg
43	19	10.0	https://pbs.twimg.com/media/CUJPNjOWsAAZRqP.jpg
46	19	9.0	https://pbs.twimg.com/media/CUJppKJWoAA75NP.jpg
62	22	11.0	https://pbs.twimg.com/media/CUY60usWoAADBxx.jpg
76	16	10.0	https://pbs.twimg.com/media/CT9Vn7PW0AA_ZCM.jpg
78	16	10.0	https://pbs.twimg.com/media/CT5Vg_wXIAAXfnj.jpg
94	16	9.0	https://pbs.twimg.com/media/CT5PY90WoAAQGLo.jpg
110	18	13.0	https://pbs.twimg.com/media/DCnll_dUQAAkBdG.jpg
128	22	13.0	https://pbs.twimg.com/media/DFWra-3VYAA2piG.jpg
141	23	12.0	https://pbs.twimg.com/media/DFYRgsOUQAARGh0.jpg
150	29	12.0	https://pbs.twimg.com/media/DDgK-J4XUAIEV9W.jpg
154	4	13.0	https://pbs.twimg.com/media/DD2oC12WAAEI_4a.jpg
155	4	13.0	https://pbs.twimg.com/media/DD5yKdPW0AArZx8.jpg
158	7	13.0	https://pbs.twimg.com/media/DEF2-_hXoAAs62q.jpg
159	8	13.5	https://pbs.twimg.com/media/DELC9dZXUADqUk.jpg
170	23	13.0	https://pbs.twimg.com/media/CZX8nyeVAAEstKM.jpg
171	23	11.0	https://pbs.twimg.com/media/CZbIIM-WkAIPClg.jpg
185	28	12.0	https://pbs.twimg.com/media/CZx_wV2UMAArgsJ.jpg
257	5	11.0	https://pbs.twimg.com/media/CX7br3HWsAAQ9L1.jpg
262	2	12.0	https://pbs.twimg.com/media/CXrawAhWkAAWSxC.jpg
294	16	12.0	https://pbs.twimg.com/media/CY2iwGNWUAI5zWi.jpg
332	27	11.0	https://pbs.twimg.com/media/CehIzzZWQAEyHH5.jpg
344	13	10.0	https://pbs.twimg.com/media/CdYzWuYUIAAHPkB.jpg
354	15	12.0	https://pbs.twimg.com/media/CdjaSFCWAAAjZh3.jpg
370	10	10.0	https://pbs.twimg.com/media/CfpNGTHUIAAA8XC.jpg
372	11	10.0	https://pbs.twimg.com/media/Cfuba6NW4AIeMHk.jpg
385	23	10.0	https://pbs.twimg.com/media/CgsA5eFWgAAu0qn.jpg
398	31	12.0	https://pbs.twimg.com/media/Ce6b4MPWwAA22Xm.jpg
...
1376	28	13.0	https://pbs.twimg.com/media/C79sB4xXwAEvwKY.jpg
1381	23	13.0	https://pbs.twimg.com/media/C5UAqgyXAAAbMWH.jpg
1382	12	10.0	https://pbs.twimg.com/media/Cf3sH62VAAA-LiP.jpg
1391	2	11.0	https://pbs.twimg.com/media/CfAHv83UMAIEQYx.jpg
1393	7	13.0	https://pbs.twimg.com/ext_tw_video_thumb/79546...
1395	13	13.0	https://pbs.twimg.com/media/Cum5LlfWAAAYPcS.jpg
1403	3	11.0	https://pbs.twimg.com/media/CtZgXgeXYAA1Gxw.jpg
1418	1	13.0	https://pbs.twimg.com/media/C-s5oYZXkAAMHHq.jpg

1419	18	13.0	https://pbs.twimg.com/media/C2d_vnHWEAE9phX.jpg
1429	20	12.0	https://pbs.twimg.com/media/CxuM3oZW8AEh05z.jpg
1443	25	10.0	https://pbs.twimg.com/media/Cg2bKLAWwAAOWEm.jpg
1479	19	12.0	https://pbs.twimg.com/media/CWnznDTU4AAa-6P.jpg
1481	14	11.0	https://pbs.twimg.com/media/CWN0dIpWoAAWid2.jpg
1498	24	11.0	https://pbs.twimg.com/media/CW-loUBWYAAAn2Cb.jpg
1522	12	12.0	https://pbs.twimg.com/media/Czb4iFRXgAIUMiN.jpg
1524	10	14.0	https://pbs.twimg.com/media/CzU_YVGUUA3Xsd.jpg
1535	8	12.0	https://pbs.twimg.com/ext_tw_video_thumb/75145...
1537	17	11.0	https://pbs.twimg.com/media/CbYyCMcWIAHHjF.jpg
1540	16	12.0	https://pbs.twimg.com/media/CbTj--1XEAIZjc_.jpg
1541	6	8.0	https://pbs.twimg.com/media/CVgGc9hWIAIe1bn.jpg
1551	18	10.0	https://pbs.twimg.com/media/CbdpBmLUYAY9SgQ.jpg
1553	4	12.0	https://pbs.twimg.com/media/CVWicBbUYAIomjC.jpg
1573	23	12.0	https://pbs.twimg.com/media/CoEwMXeWEAAaIz5.jpg
1579	4	12.0	https://pbs.twimg.com/media/C6CI_jbVAAA3-a1.jpg
1584	16	11.0	https://pbs.twimg.com/media/Cp8k6oRWcAUL78U.jpg
1585	30	11.0	https://pbs.twimg.com/media/CZ724fDUYAAytS-.jpg
1588	12	12.0	https://pbs.twimg.com/ext_tw_video_thumb/67574...
1598	31	12.0	https://pbs.twimg.com/media/CotUFZEwCAA2Pku.jpg
1615	3	13.0	https://pbs.twimg.com/media/C-3wvtXxXcAUTuBE.jpg
1617	24	13.0	https://pbs.twimg.com/media/DFg_2PVW0AEHN3p.jpg

	dog_type	confidence_list	favorite_count	retweet_count
8	golden_retriever	0.633037	476.0	245.0
30	golden_retriever	0.765266	186.0	51.0
37	golden_retriever	0.471351	260.0	62.0
43	golden_retriever	0.949892	14446.0	6300.0
46	golden_retriever	0.462556	497.0	245.0
62	golden_retriever	0.215944	517.0	142.0
76	golden_retriever	0.858744	290.0	133.0
78	golden_retriever	0.775930	471.0	215.0
94	golden_retriever	0.007959	294.0	141.0
110	golden_retriever	0.874566	18521.0	2371.0
128	golden_retriever	0.469760	25188.0	4224.0
141	golden_retriever	0.714719	28667.0	4422.0
150	golden_retriever	0.913255	28137.0	6188.0
154	golden_retriever	0.006941	28952.0	4903.0
155	golden_retriever	0.762211	44332.0	11434.0
158	golden_retriever	0.949562	36675.0	6575.0
159	golden_retriever	0.943082	45272.0	9803.0
170	golden_retriever	0.883229	3987.0	1372.0
171	golden_retriever	0.862964	3782.0	1345.0
185	golden_retriever	0.636845	4578.0	1648.0
257	golden_retriever	0.203249	1308.0	229.0
262	golden_retriever	0.889848	2286.0	709.0
294	golden_retriever	0.900437	9791.0	4781.0
332	golden_retriever	0.371816	2971.0	793.0

344	golden_retriever	0.976139	17320.0	7433.0
354	golden_retriever	0.790026	3453.0	1136.0
370	golden_retriever	0.140394	3659.0	1169.0
372	golden_retriever	0.765778	4631.0	1335.0
385	golden_retriever	0.839390	3171.0	951.0
398	golden_retriever	0.990715	4536.0	1730.0
...
1376	golden_retriever	0.650003	46517.0	12341.0
1381	golden_retriever	0.226564	14452.0	2737.0
1382	golden_retriever	0.605304	5060.0	1862.0
1391	golden_retriever	0.638625	7039.0	2049.0
1393	golden_retriever	0.193082	53637.0	26326.0
1395	golden_retriever	0.929266	11722.0	3864.0
1403	golden_retriever	0.560699	8016.0	1535.0
1418	golden_retriever	0.578120	15740.0	3566.0
1419	golden_retriever	0.980071	8986.0	1821.0
1429	golden_retriever	0.359860	11999.0	3103.0
1443	golden_retriever	0.240695	5557.0	1758.0
1479	golden_retriever	0.853284	1807.0	573.0
1481	golden_retriever	0.790386	12073.0	4950.0
1498	golden_retriever	0.698961	3757.0	1467.0
1522	golden_retriever	0.426183	9362.0	2421.0
1524	golden_retriever	0.873233	15649.0	4061.0
1535	golden_retriever	0.714409	3395.0	1080.0
1537	golden_retriever	0.808978	3217.0	1029.0
1540	golden_retriever	0.836572	2285.0	543.0
1541	golden_retriever	0.889241	7687.0	3286.0
1551	golden_retriever	0.532460	7938.0	2950.0
1553	golden_retriever	0.952773	1321.0	513.0
1573	golden_retriever	0.790371	7065.0	2180.0
1579	golden_retriever	0.887625	35974.0	8466.0
1584	golden_retriever	0.829456	7562.0	2350.0
1585	golden_retriever	0.989333	2351.0	536.0
1588	golden_retriever	0.800495	1210.0	366.0
1598	golden_retriever	0.985876	6337.0	2065.0
1615	golden_retriever	0.895529	18804.0	1612.0
1617	golden_retriever	0.953442	14884.0	2208.0

[142 rows x 5 columns]

1.3 Insight two

Japanese_spaniel has the lowest average rating Clumber has the highest average rating

```
In [12]: df_dog_type_mean = df_master.groupby('dog_type').mean()
df_dog_type_sorted = df_dog_type_mean['rating'].sort_values()

df_dog_type_sorted
```

```

Out[12]: dog_type
Japanese_spaniel          5.000000
Afghan_hound              8.000000
soft-coated_wheaten_terrier 8.800000
Walker_hound              9.000000
Scotch_terrier            9.000000
Tibetan_terrier           9.250000
dalmatian                 9.272727
Boston_bull               9.416667
Saint_Bernard             9.500000
Welsh_springer_spaniel    9.500000
Dandie_Dinmont           9.571429
Norwich_terrier           9.600000
miniature_schnauzer       9.600000
redbone                   9.666667
English_setter            9.714286
Maltese_dog               9.736842
Rhodesian_ridgeback       9.750000
Scottish_deerhound        9.750000
Irish_terrier             9.800000
Airedale                  9.833333
papillon                  9.833333
Newfoundland              9.857143
Mexican_hairless          9.857143
miniature_poodle          9.875000
miniature_pinscher        9.916667
Brabancon_griffon         10.000000
Italian_greyhound         10.000000
groenendael               10.000000
Ibizan_hound              10.000000
beagle                    10.052632

...

Rottweiler                11.105263
cocker_spaniel            11.153846
Labrador_retriever        11.194175
clumber                   11.270000
Bernese_mountain_dog      11.272727
kelpie                    11.307692
Greater_Swiss_Mountain_dog 11.333333
Irish_water_spaniel       11.333333
Leonberg                 11.333333
Doberman                  11.333333
chow                      11.377778
Pembroke                  11.393617
Eskimo_dog                11.409091
Siberian_husky            11.425000
Norfolk_terrier           11.428571
Great_Pyrenees            11.466667

```

Australian_terrier	11.500000
wire-haired_fox_terrier	11.500000
giant_schnauzer	11.500000
golden_retriever	11.588028
Samoyed	11.657895
Gordon_setter	11.750000
silky_terrier	12.000000
standard_schnauzer	12.000000
Tibetan_mastiff	12.000000
Border_terrier	12.142857
Irish_setter	12.333333
briard	12.333333
Saluki	12.500000
Bouvier_des_Flandres	13.000000

Name: rating, Length: 113, dtype: float64

```
In [15]: print(df_master.loc[df_master.dog_type == 'Japanese_spaniel', 'jpg_url'])
```

```
710      https://pbs.twimg.com/media/CWZBsJPWsAAZFF1.jpg
Name: jpg_url, dtype: object
```

Please check the url, this Japanese Spaniel has the lowest average rating. Poor guy!
Here are examples one of the the highest rated dog types. The Briard:

```
In [18]: print(df_master.loc[df_master.dog_type == 'briard', 'jpg_url'])
```

```
216      https://pbs.twimg.com/media/CapsyfkWcAQ41uC.jpg
1401     https://pbs.twimg.com/media/CttPBtOWIAAcSDE.jpg
1533     https://pbs.twimg.com/media/C41st0bXAAE6MP8.jpg
Name: jpg_url, dtype: object
```

1.4 Insight three & visualization

Dog_types with low number of ratings show a high variaty of mean ratings

```
In [20]: df_dog_type_count = df_master.groupby('dog_type').count()
df_dog_type_count
```

```
Out[20]:
```

	tweet_id	text	rating_numerator	\
dog_type				
Afghan_hound	2	2	2	
Airedale	12	12	12	
American_Staffordshire_terrier	16	16	16	
Appenzeller	2	2	2	
Australian_terrier	2	2	2	
Bedlington_terrier	6	6	6	
Bernese_mountain_dog	11	11	11	

Blenheim_spaniel	10	10	10
Border_collie	12	12	12
Border_terrier	7	7	7
Boston_bull	12	12	12
Bouvier_des_Flandres	1	1	1
Brabancon_griffon	3	3	3
Brittany_spaniel	8	8	8
Cardigan	19	19	19
Chesapeake_Bay_retriever	31	31	31
Chihuahua	87	87	87
Dandie_Dinmont	7	7	7
Doberman	9	9	9
English_setter	7	7	7
English_springer	9	9	9
EntleBucher	1	1	1
Eskimo_dog	22	22	22
French_bulldog	30	30	30
German_shepherd	21	21	21
German_short-haired_pointer	8	8	8
Gordon_setter	4	4	4
Great_Dane	11	11	11
Great_Pyrenees	15	15	15
Greater_Swiss_Mountain_dog	3	3	3
...
cocker_spaniel	26	26	26
collie	11	11	11
curly-coated_retriever	3	3	3
dalmatian	11	11	11
flat-coated_retriever	8	8	8
giant_schnauzer	4	4	4
golden_retriever	142	142	142
groenendael	2	2	2
keeshond	4	4	4
kelpie	13	13	13
komondor	3	3	3
kuvasz	18	18	18
malamute	32	32	32
malinois	11	11	11
miniature_pinscher	24	24	24
miniature_poodle	8	8	8
miniature_schnauzer	5	5	5
papillon	6	6	6
pug	60	60	60
redbone	6	6	6
schipperke	11	11	11
silky_terrier	1	1	1
soft-coated_wheaten_terrier	15	15	15
standard_poodle	10	10	10

standard_schnauzer	1	1	1
toy_poodle	50	50	50
toy_terrier	3	3	3
vizsla	13	13	13
whippet	11	11	11
wire-haired_fox_terrier	2	2	2

	rating_denominator	name	dog_stage	year	\
dog_type					
Afghan_hound	2	2	2	2	
Airedale	12	12	12	12	
American_Staffordshire_terrier	16	16	16	16	
Appenzeller	2	2	2	2	
Australian_terrier	2	2	2	2	
Bedlington_terrier	6	6	6	6	
Bernese_mountain_dog	11	11	11	11	
Blenheim_spaniel	10	10	10	10	
Border_collie	12	12	12	12	
Border_terrier	7	7	7	7	
Boston_bull	12	12	12	12	
Bouvier_des_Flandres	1	1	1	1	
Brabancon_griffon	3	3	3	3	
Brittany_spaniel	8	8	8	8	
Cardigan	19	19	19	19	
Chesapeake_Bay_retriever	31	31	31	31	
Chihuahua	87	87	87	87	
Dandie_Dinmont	7	7	7	7	
Doberman	9	9	9	9	
English_setter	7	7	7	7	
English_springer	9	9	9	9	
EntleBucher	1	1	1	1	
Eskimo_dog	22	22	22	22	
French_bulldog	30	30	30	30	
German_shepherd	21	21	21	21	
German_short-haired_pointer	8	8	8	8	
Gordon_setter	4	4	4	4	
Great_Dane	11	11	11	11	
Great_Pyrenees	15	15	15	15	
Greater_Swiss_Mountain_dog	3	3	3	3	
...	
cocker_spaniel	26	26	26	26	
collie	11	11	11	11	
curly-coated_retriever	3	3	3	3	
dalmatian	11	11	11	11	
flat-coated_retriever	8	8	8	8	
giant_schnauzer	4	4	4	4	
golden_retriever	142	142	142	142	
groenendael	2	2	2	2	

keeshond	4	4	4	4
kelpie	13	13	13	13
komondor	3	3	3	3
kuvasz	18	18	18	18
malamute	32	32	32	32
malinois	11	11	11	11
miniature_pinscher	24	24	24	24
miniature_poodle	8	8	8	8
miniature_schnauzer	5	5	5	5
papillon	6	6	6	6
pug	60	60	60	60
redbone	6	6	6	6
schipperke	11	11	11	11
silky_terrier	1	1	1	1
soft-coated_wheaten_terrier	15	15	15	15
standard_poodle	10	10	10	10
standard_schnauzer	1	1	1	1
toy_poodle	50	50	50	50
toy_terrier	3	3	3	3
vizsla	13	13	13	13
whippet	11	11	11	11
wire-haired_fox_terrier	2	2	2	2

	month	day	rating	jpg_url	confidence_list \
dog_type					
Afghan_hound	2	2	2	2	2
Airedale	12	12	12	12	12
American_Staffordshire_terrier	16	16	16	16	16
Appenzeller	2	2	2	2	2
Australian_terrier	2	2	2	2	2
Bedlington_terrier	6	6	6	6	6
Bernese_mountain_dog	11	11	11	11	11
Blenheim_spaniel	10	10	10	10	10
Border_collie	12	12	12	12	12
Border_terrier	7	7	7	7	7
Boston_bull	12	12	12	12	12
Bouvier_des_Flandres	1	1	1	1	1
Brabancon_griffon	3	3	3	3	3
Brittany_spaniel	8	8	8	8	8
Cardigan	19	19	19	19	19
Chesapeake_Bay_retriever	31	31	31	31	31
Chihuahua	87	87	87	87	87
Dandie_Dinmont	7	7	7	7	7
Doberman	9	9	9	9	9
English_setter	7	7	7	7	7
English_springer	9	9	9	9	9
EntleBucher	1	1	1	1	1
Eskimo_dog	22	22	22	22	22

French_bulldog	30	30	30	30	30
German_shepherd	21	21	21	21	21
German_short-haired_pointer	8	8	8	8	8
Gordon_setter	4	4	4	4	4
Great_Dane	11	11	11	11	11
Great_Pyrenees	15	15	15	15	15
Greater_Swiss_Mountain_dog	3	3	3	3	3
...
cocker_spaniel	26	26	26	26	26
collie	11	11	11	11	11
curly-coated_retriever	3	3	3	3	3
dalmatian	11	11	11	11	11
flat-coated_retriever	8	8	8	8	8
giant_schnauzer	4	4	4	4	4
golden_retriever	142	142	142	142	142
groenendael	2	2	2	2	2
keeshond	4	4	4	4	4
kelpie	13	13	13	13	13
komondor	3	3	3	3	3
kuvasz	18	18	18	18	18
malamute	32	32	32	32	32
malinois	11	11	11	11	11
miniature_pinscher	24	24	24	24	24
miniature_poodle	8	8	8	8	8
miniature_schnauzer	5	5	5	5	5
papillon	6	6	6	6	6
pug	60	60	60	60	60
redbone	6	6	6	6	6
schipperke	11	11	11	11	11
silky_terrier	1	1	1	1	1
soft-coated_wheaten_terrier	15	15	15	15	15
standard_poodle	10	10	10	10	10
standard_schnauzer	1	1	1	1	1
toy_poodle	50	50	50	50	50
toy_terrier	3	3	3	3	3
vizsla	13	13	13	13	13
whippet	11	11	11	11	11
wire-haired_fox_terrier	2	2	2	2	2

	favorite_count	retweet_count
dog_type		
Afghan_hound	2	2
Airedale	12	12
American_Staffordshire_terrier	16	16
Appenzeller	2	2
Australian_terrier	2	2
Bedlington_terrier	6	6
Bernese_mountain_dog	11	11

Blenheim_spaniel	10	10
Border_collie	12	12
Border_terrier	7	7
Boston_bull	12	12
Bouvier_des_Flandres	1	1
Brabancon_griffon	3	3
Brittany_spaniel	8	8
Cardigan	19	19
Chesapeake_Bay_retriever	31	31
Chihuahua	87	87
Dandie_Dinmont	7	7
Doberman	9	9
English_setter	7	7
English_springer	9	9
EntleBucher	1	1
Eskimo_dog	22	22
French_bulldog	29	29
German_shepherd	21	21
German_short-haired_pointer	8	8
Gordon_setter	4	4
Great_Dane	11	11
Great_Pyrenees	15	15
Greater_Swiss_Mountain_dog	3	3
...
cocker_spaniel	26	26
collie	11	11
curly-coated_retriever	3	3
dalmatian	11	11
flat-coated_retriever	8	8
giant_schnauzer	4	4
golden_retriever	142	142
groenendael	2	2
keeshond	4	4
kelpie	13	13
komondor	3	3
kuvasz	18	18
malamute	32	32
malinois	11	11
miniature_pinscher	24	24
miniature_poodle	8	8
miniature_schnauzer	5	5
papillon	6	6
pug	60	60
redbone	6	6
schipperke	11	11
silky_terrier	1	1
soft-coated_wheaten_terrier	15	15
standard_poodle	10	10

standard_schnauzer	1	1
toy_poodle	50	50
toy_terrier	3	3
vizsla	13	13
whippet	11	11
wire-haired_fox_terrier	2	2

[113 rows x 14 columns]

```
In [21]: dog_type_count = df_dog_type_count['rating']
dog_type_mean = df_dog_type_mean['rating']
dog_type_mean
```

```
Out[21]: dog_type
Afghan_hound          8.000000
Airedale              9.833333
American_Staffordshire_terrier  10.312500
Appenzeller          11.000000
Australian_terrier    11.500000
Bedlington_terrier    10.333333
Bernese_mountain_dog  11.272727
Blenheim_spaniel      11.100000
Border_collie         10.583333
Border_terrier        12.142857
Boston_bull           9.416667
Bouvier_des_Flandres  13.000000
Brabancon_griffon     10.000000
Brittany_spaniel      10.375000
Cardigan              11.105263
Chesapeake_Bay_retriever  10.838710
Chihuahua             10.448276
Dandie_Dinmont        9.571429
Doberman              11.333333
English_setter         9.714286
English_springer      10.555556
EntleBucher           11.000000
Eskimo_dog            11.409091
French_bulldog        11.100000
German_shepherd       11.000000
German_short-haired_pointer  10.500000
Gordon_setter         11.750000
Great_Dane            10.818182
Great_Pyrenees        11.466667
Greater_Swiss_Mountain_dog  11.333333
...
cocker_spaniel        11.153846
collie                10.909091
curly-coated_retriever  10.333333
```

dalmatian	9.272727
flat-coated_retriever	10.875000
giant_schnauzer	11.500000
golden_retriever	11.588028
groenendael	10.000000
keeshond	10.250000
kelpie	11.307692
komondor	11.000000
kuvasz	10.792222
malamute	10.875000
malinois	10.818182
miniature_pinscher	9.916667
miniature_poodle	9.875000
miniature_schnauzer	9.600000
papillon	9.833333
pug	10.200000
redbone	9.666667
schipperke	10.818182
silky_terrier	12.000000
soft-coated_wheaten_terrier	8.800000
standard_poodle	10.400000
standard_schnauzer	12.000000
toy_poodle	11.020000
toy_terrier	10.333333
vizsla	10.846154
whippet	10.181818
wire-haired_fox_terrier	11.500000

Name: rating, Length: 113, dtype: float64

```
In [22]: df = pd.DataFrame()
df['dog_type_count'] = dog_type_count
df['dog_type_mean'] = dog_type_mean
df
```

```
Out[22]:
```

	dog_type_count	dog_type_mean
dog_type		
Afghan_hound	2	8.000000
Airedale	12	9.833333
American_Staffordshire_terrier	16	10.312500
Appenzeller	2	11.000000
Australian_terrier	2	11.500000
Bedlington_terrier	6	10.333333
Bernese_mountain_dog	11	11.272727
Blenheim_spaniel	10	11.100000
Border_collie	12	10.583333
Border_terrier	7	12.142857
Boston_bull	12	9.416667
Bouvier_des_Flandres	1	13.000000

Brabancon_griffon	3	10.000000
Brittany_spaniel	8	10.375000
Cardigan	19	11.105263
Chesapeake_Bay_retriever	31	10.838710
Chihuahua	87	10.448276
Dandie_Dinmont	7	9.571429
Doberman	9	11.333333
English_setter	7	9.714286
English_springer	9	10.555556
EntleBucher	1	11.000000
Eskimo_dog	22	11.409091
French_bulldog	30	11.100000
German_shepherd	21	11.000000
German_short-haired_pointer	8	10.500000
Gordon_setter	4	11.750000
Great_Dane	11	10.818182
Great_Pyrenees	15	11.466667
Greater_Swiss_Mountain_dog	3	11.333333
...
cocker_spaniel	26	11.153846
collie	11	10.909091
curly-coated_retriever	3	10.333333
dalmatian	11	9.272727
flat-coated_retriever	8	10.875000
giant_schnauzer	4	11.500000
golden_retriever	142	11.588028
groenendael	2	10.000000
keeshond	4	10.250000
kelpie	13	11.307692
komondor	3	11.000000
kuvasz	18	10.792222
malamute	32	10.875000
malinois	11	10.818182
miniature_pinscher	24	9.916667
miniature_poodle	8	9.875000
miniature_schnauzer	5	9.600000
papillon	6	9.833333
pug	60	10.200000
redbone	6	9.666667
schipperke	11	10.818182
silky_terrier	1	12.000000
soft-coated_wheaten_terrier	15	8.800000
standard_poodle	10	10.400000
standard_schnauzer	1	12.000000
toy_poodle	50	11.020000
toy_terrier	3	10.333333
vizsla	13	10.846154
whippet	11	10.181818

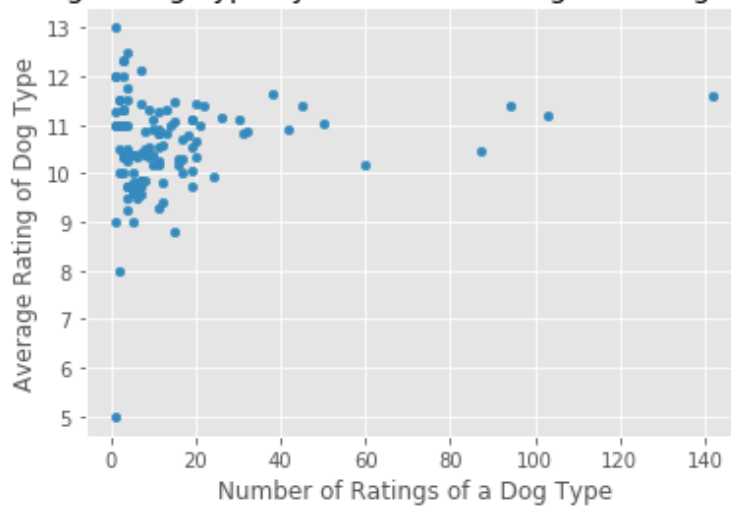
```
wire-haired_fox_terrier                2          11.500000
```

```
[113 rows x 2 columns]
```

```
In [23]: df.plot(x='dog_type_count', y='dog_type_mean', kind='scatter')
plt.xlabel('Number of Ratings of a Dog Type')
plt.ylabel('Average Rating of Dog Type')
plt.title('Average Rating of Dog Type by Number of Ratings of a Dog Type Scatter Plot')

fig = plt.gcf()
#plt.savefig('X:/' + newName + '.png',
fig.savefig('output2.png',bbox_inches='tight');
```

Average Rating of Dog Type by Number of Ratings of a Dog Type Scatter Plot



The blue dot on the very right represents the approximately 140 ratings of golden retriever. On average these ratings are high 11.59. This really shows that the golden retriever is very popular.

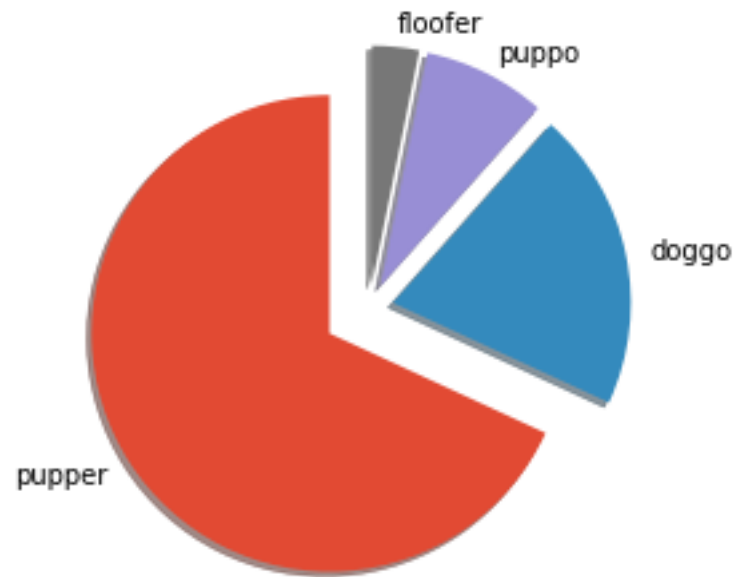
1.5 Insight four & visualization

Are there more male or female dogs?

```
In [25]: dog_stage_count = list(df_master[df_master['dog_stage'] != 'None']['dog_stage'].value_counts())
dog_stages = df_master[df_master['dog_stage'] != 'None']['dog_stage'].value_counts().index
explode = (0.2, 0.1, 0.1, 0.1)
```

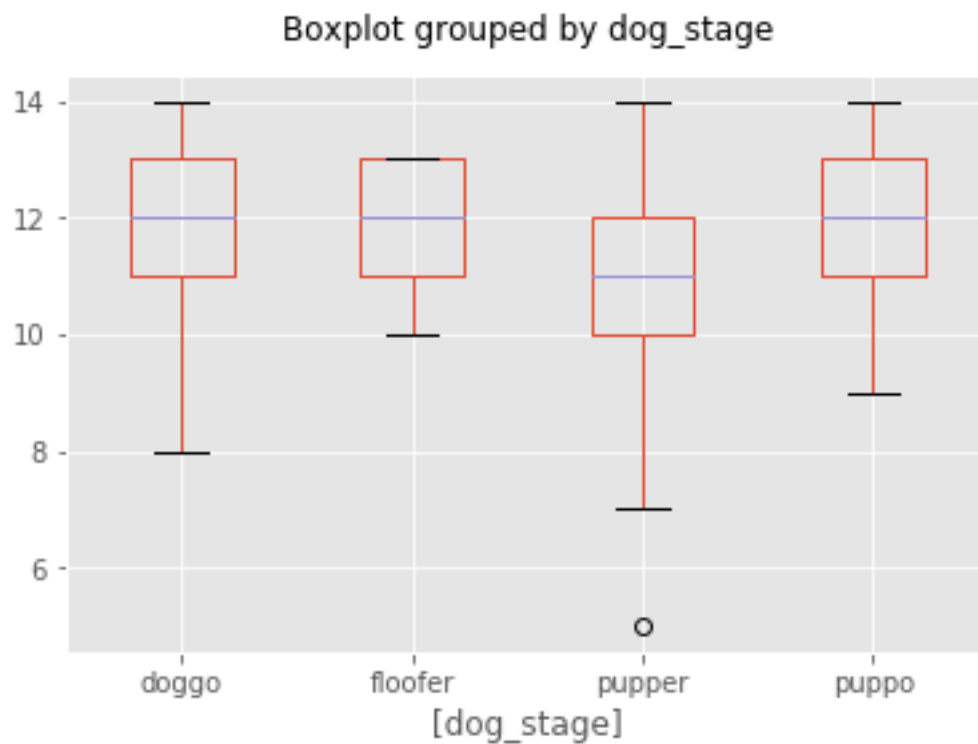
```
fig1, ax1 = plt.subplots()
ax1.pie(dog_stage_count, explode = explode, labels = dog_stages, shadow = True, startangle=90)
ax1.axis('equal')
```

```
Out [25]: (-1.2887489155281924,
1.2353554618963543,
-1.2286790096332167,
1.2103647382990086)
```



```
In [26]: df_master[df_master['dog_stage'] != 'None'].boxplot(column = ['rating_numerator'], by  
plt.title('')
```

```
Out[26]: Text(0.5, 1.0, '')
```



```
In [27]: df_master[df_master['dog_stage'] != 'None'].groupby('dog_stage')['rating_numerator'].r
```

```
Out[27]: dog_stage
doggo      11.862745
floofer    11.875000
pupper     10.837836
puppo      12.000000
Name: rating_numerator, dtype: float64
```

Puppies represent the big number of our pie, but it has the lowest mean rating

1.6 Conclusion

The Twitter account WeRateDogs (@dog_rates) is devoted to humorously reviewing pictures of dogs doing adorable poses. Dogs are rated on a scale of one to ten, but are invariably given ratings in excess of the maximum, such as "13/10". It has acquired over 4.50 million followers since its debut.