

## Analysis of WeRateDogs from Twitter

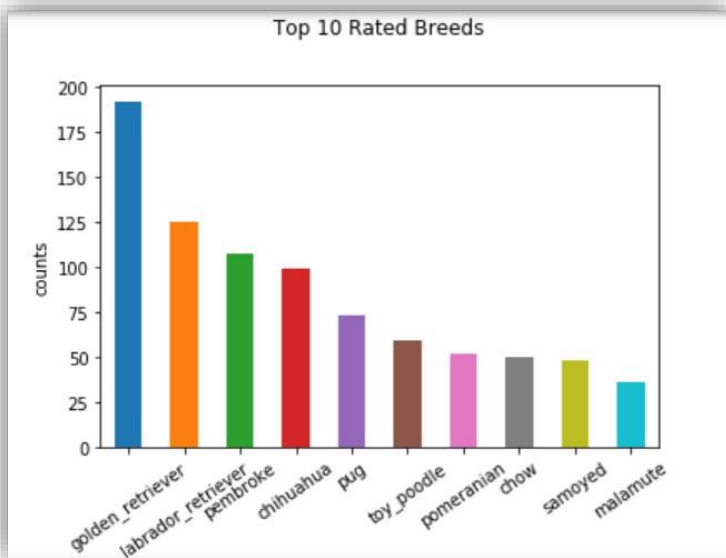
This report is to wrangle, analyze, and visualize the tweet archive of Twitter user @dog\_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

After wrangling the data, fixing quality issues and structural issues, we only use non-retweets for this analysis. We found these insights below.

What are the top 10 rated dog breeds?

The top 10 dog breeds are Golden retriever, Labrador retriever, Pembroke, Chihuahua, Pug, Toy Poodle, Pomeranian, Chow, Samoyed, and Malamute. Golden retriever is the top 1 and it is much higher than the second one Labrador retriever.

Here is the web site explaining why people love golden retriever so much and why they are the super dog: <https://mypuppystory.com/25-reasons-why-golden-retrievers-are-superior-dogs/>



What are the top 5 most common dog names?

Top 1 is **Charlie**, top 2 is **Oliver**, top 3 is **Cooper**, top 4 is **Lucy**, top 5 is **Penny** (None is ignored). However, most of the dog owners don't properly show their dogs' names. There are 639 "None" values.

What are the correlation between rating and favorite count for different dog stage?

The definition of rating here is rating\_numerator divided by rating\_denominator.

The range of the rating is generally between 0 to 2.7, but there is an outlier 177.6 in None group. Most of them (25%-75%) are between 1 to 1.3.

	count	mean	std	min	25%	50%	75%	max
dog_stage								
None	1688.0	1.180085	4.419681	0.0	1.000	1.1	1.2	177.6
doggo	63.0	1.188889	0.147135	0.8	1.100	1.2	1.3	1.4
doggo,floofer	1.0	1.100000	NaN	1.1	1.100	1.1	1.1	1.1
doggo,pupper	9.0	1.100000	0.239792	0.5	1.100	1.2	1.2	1.3
doggo,puppo	1.0	1.300000	NaN	1.3	1.300	1.3	1.3	1.3
floofer	7.0	1.200000	0.115470	1.0	1.150	1.2	1.3	1.3
pupper	203.0	1.071429	0.207463	0.3	1.000	1.1	1.2	2.7
puppo	22.0	1.200000	0.130931	0.9	1.125	1.2	1.3	1.4

Most of the data are below 20000 favorite counts and rating near 1. Red dots stand for pupper have the largest counts among dog stages except None. Many of pupper don't have high favorite counts (less than 5000). We also notice that doggos (blue dots) tend to have higher values compared to the pupper (red dots) which tend to have more lower values than higher values.

