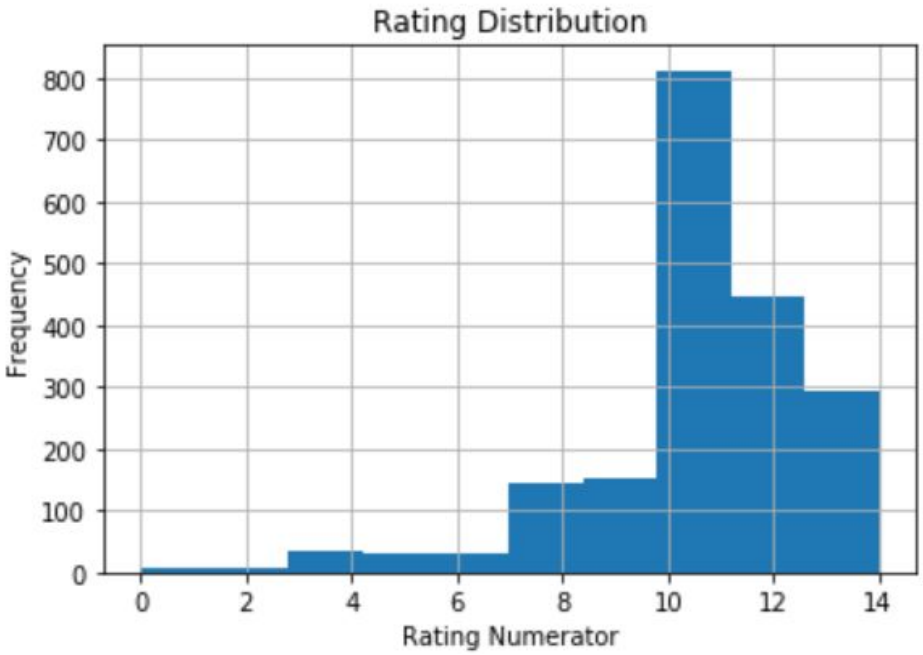
Hacking the WeRateDogs Rating System

By Patrick Maloney

After digging into the tweet archive of the WeRateDogs Twitter feed, I looked for some insights to see if there was a way to figure out how to score a higher rating. Data was collected from an archive of tweets, the twitter API, and a neural network algorithm that helped us analyze the photos to predict the breed of each dog.

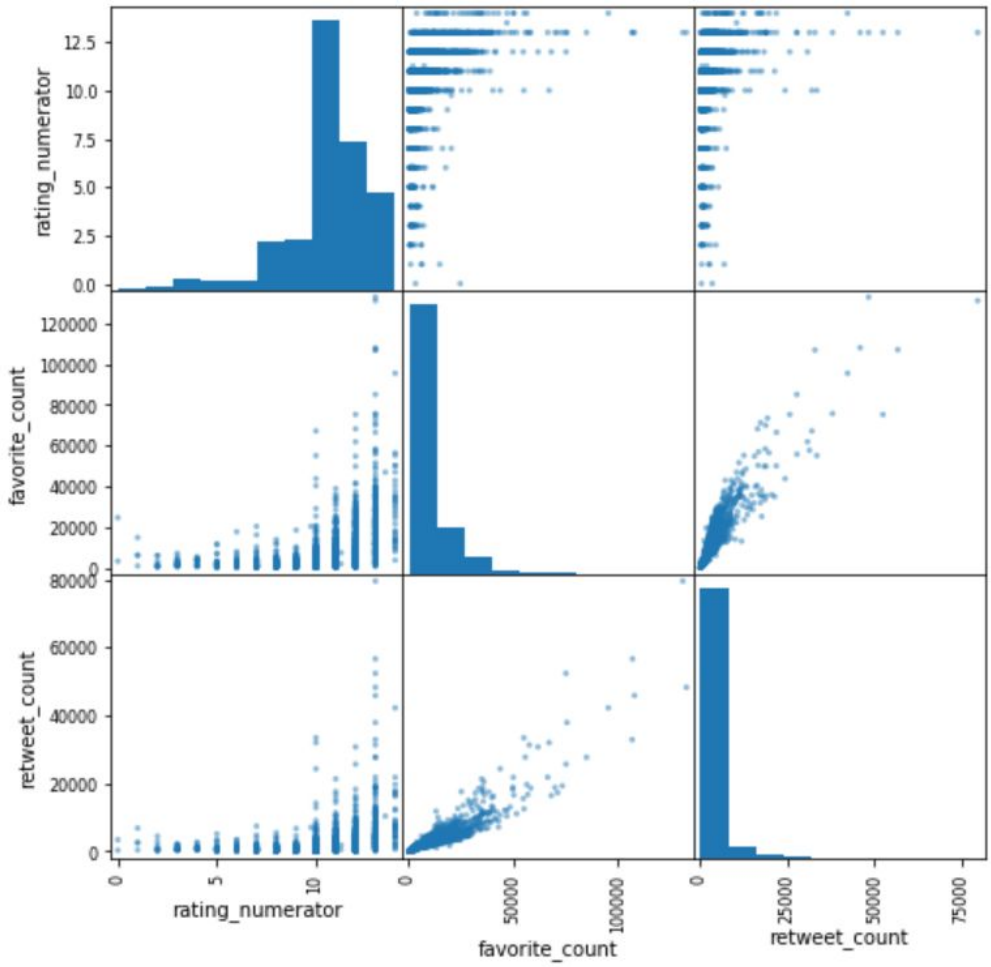
The most common breed in the data set, unsurprisingly, was the catch-all 'other' category, which can contain mutts/mixed breeds, as well as all others that the prediction algorithm was unable to predict. Aside from these, golden retriever were the breed most submitted for ratings, followed by labradors, then pembrokes, then Chihuahuas.



The average rating is 10.55, so you want to get at least an 11 if you're hoping your dog stands out among the rest of the pack. If you got a 12, that put you in the top 37%, while a 13 would put you in the top 15%, and if you manage to snag a rating of 14, pat your dog on the back, because you can

go to sleep knowing that your dog is rate higher than more than 98% of the dogs submitted. Give him a treat!

The rating system is crazy, but still very popular. Let's see how the rest of the Twitterverse feels about the opinions of @WeRateDogs.



Interestingly, the scatter matrix above suggests that Twitter seems to agree with the rating system, since higher-rated dogs appear to get more favorites and retweets. Unsurprisingly, there is a strong positive correlation between favorites and retweets: the more a dog is retweeted, the more favorites it gets, while the more times a dog is favorited, it is more likely to be retweeted, which in turn makes it more likely to get more favorites.

What if I want to go buy a dog for the sole purpose of scoring a high rating on @WeRateDogs? Are there certain traits I can seek out that will help my chances? Let's look at the dog_stages to see if we can find anything out.

Dep. Variable:	rating_numerator	**R-squared:**	0.020
Model:	OLS	**Adj. R-squared:**	0.018
Method:	Least Squares	**F-statistic:**	8.228
Date:	Fri, 22 Feb 2019	**Prob (F-statistic):**	1.06e-07
Time:	14:33:02	**Log-Likelihood:**	-4319.6
No. Observations:	1974	**AIC:**	8651.
Df Residuals:	1968	**BIC:**	8685.
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
intercept	10.6368	0.152	70.111	0.000	10.339	10.934
doggo	1.2521	0.312	4.016	0.000	0.641	1.863
floofer	1.3632	0.831	1.640	0.101	-0.266	2.993
pupper	1.3632	0.485	2.810	0.005	0.412	2.315
multiple	0.5450	0.669	0.814	0.415	-0.767	1.857
puppo	-0.1731	0.161	-1.077	0.282	-0.488	0.142

Omnibus:	550.284	**Durbin-Watson:**	1.502
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	1466.987
Skew:	-1.470	**Prob(JB):**	0.00
Kurtosis:	6.032	**Cond. No.**	22.8

In this regression model, dogs with no dog_stage value were used as the baseline. The model suggests that having a dog in one of these stages will likely increase your rating (as long as it isn't a puppo), including by a whopping +1.36 points if your dog is deemed a "floofer" or a "pupper" Having a "doggo" could improve your rating by +1.25, and having multiple stages may improve your score by half a point. The model suggests that there may be a bit less enthusiasm for "puppos," which can lower your rating by -0.17 points. So why does @WeRateDogs dislike puppos? Well, chances are that isn't actually the case, as this variable's p-value suggests that this particular coefficient is likely irrelevant. The same goes for floofers.

So to get a high rating, I should get a floofer-pupper and I'll get a high rating, right? Not necessarily. The R-squared value suggests a high level of covariance, and means it was very difficult to fit this model. With an R-squared value of 0.02, we can have some fun thinking about why the rater has a subconscious bias for floofers and puppos, we are probably better off taking this with a grain of salt and embracing the randomness of the ratings.