# WeRateDogs Twitter Analysis

By Nanli



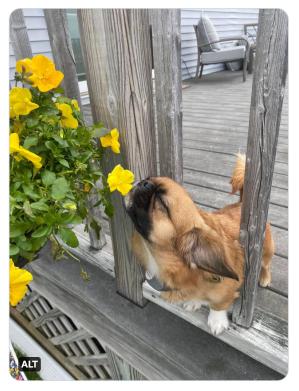
Source: https://twitter.com/dog\_rates?ref\_src=twsrc%5Eqoogle%7Ctwcamp%5Eserp%7Ctwqr%5Eauthor

If you go through a Twitter Account named WeRateDogs, you see people rate dogs with humorous comment and photo of the dog. Recently, in a Udacity Data Analyst Project, I collected different data sources about this account, played around the data by accessing, cleaning, and finally analyzing. Here I am going to share some interesting insights and visualizations I have discovered.

Before looking deeper into the fun facts about WeRateDogs, there are some information about the dogs. ©

I. In WeRateDogs tweets, people in general will introduce the dogs with their names, funny things about them, provide photos and leave a rating score for the dogs.





Source: https://twitter.com/dog\_rates?ref\_src=twsrc%5Egoogle%7Ctwcamp%5Eserp%7Ctwgr%5Eauthor

- 2. The rating system in WeRateDogs is interesting that although the rating denominator is 10, the numerators is sometimes greater than 10. For example, II/10, I2/10, I3/10, etc. And this is considered very common. Why? Because "they're good dogs Brent". Dogs are just too good, that they deserve scores more than 10!
- 3. People normally would not mention the breed of the dog, but luckily, they always leave photos. And thanks to machine leaning techniques, we can predict a dog's breed by neuro network algorithms. And this prediction values are used in the analysis.
- 4. If you're a twitter user, you know we can retweet and like the tweet, retweet count and likes are also two important variables used here.

Now that you know a bit of everything about WeRateDogs, let's look at what I have discovered.

#### Which kind of dogs get the most tweets?

I analyzed 2000+ WeRateDogs tweets and found out Golden Retriever is the most popular breed in terms of the tweets count. Well, what can I say, this is not a huge surprise, I mean who can say no to their friendly faces!

	Average Ratings	Total Tweets
Dog Breed		
golden_retriever	I2	158
labrador_retriever	II	108
pembroke	II	95
chihuahua	II	91
pug	IO	62
toy_poodle	II	<b>5</b> I
chow	II	48
pomeranian	II	42
samoyed	12	42
malamute	II	33

Golden Retriever also got a high average rating score (12) among the top 10 common breeds. Labrador Retriever and Pembroke ranked second and third, they have an average rating of II.

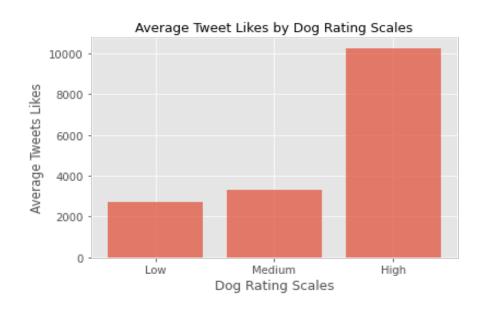
### What's the most common name for a dog?

IOOO+ names were extracted from the dataset, and I grouped the names to get the most common names for a dog. It looks like Charlie and Lucy are top I names people would like to use for their dogs. Also, the name Oliver, Cooper and Penny got quite similar counts. Dogs named Charlie there also have a high average rating of I2.

	Average Ratings	Total Tweets
Dog Name		
Charlie	I2	II
Lucy	II	II
Oliver	II	IO
Cooper	II	IO
Penny	II	9
Tucker	12	9
Winston	II	8
Lola	II	8
Sadie	IO	8
Daisy	12	7

## Who gets the most likes?

I am wondering if there is any relation between the dog ratings and the likes. Do dogs with high ratings also get more likes? The distribution for rating values and the like counts are spread out, to better visualize this topic, I grouped the ratings into different ranges. More specifically, rating scores between 0 to 5 are in the low rating scale, between 5 to 10 are in the medium rating scale, and between 10 to 15 are in the high rating scale.

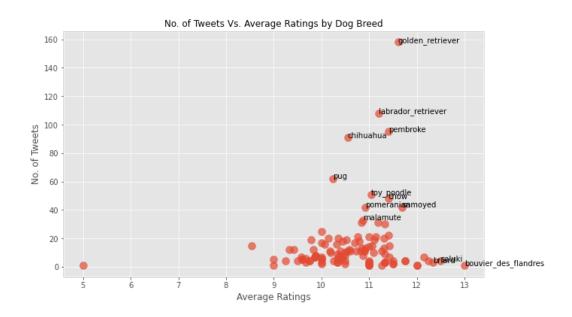


I plotted a bar chart, and it looks very clear that tweets with dogs in high rating scale have very high number of likes. Though when they are in o-5 and 5-10 ranges, the differences are small. People like the high rating dogs, these tweets attract them.

#### How spread out the Ratings and Tweets data are?

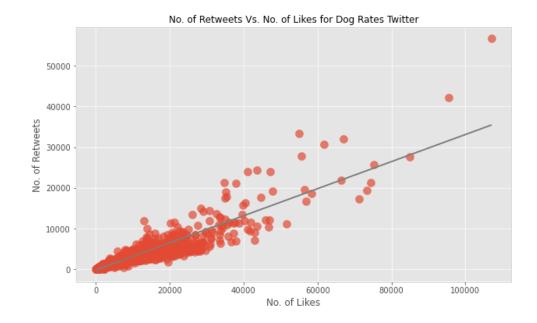
We got an idea of which dog breed has the highest number of tweets, but how about the ratings? Is an average rating of II for Golden Retriever is higher than average? Is it the highest rating at all?

While, to answer these questions, I have plotted the No. of Tweets Vs. Average Ratings by Dog Breed. The scatter plot suggests the number of tweets for Golden Retriever is much higher than average. Most of the dog ratings fall into range 10-12, so average rating of 11 is a high score. While Bouvier des Flanders dog has the highest average rating score of 13, It has a very small number of total tweets, therefore, this high rating value is less persuasive.



#### **Understand Retweets and Likes**

Retweet and like are two most common actions for tweets, I would like to understand the relation between these two actions. Do tweets have high likes also being retweeted the most?



The data is neat! After plotting the No. of Retweets Vs. No. of Likes, I can say confidently that tweets have more likes also have high number of retweets. The fitted linear regression line also proves the same!

That will be all the findings for my project, and I have to say it is a great journey to wrangle and analyze the data. Of course, I feel more than happy to share it with you guys! ©