

WeRateDogs Twitter Data from 2015 to 2017 – Report

Ask any Twitter user who loves logs, cute animals, and you'll find with near certainty that they follow *WeRateDogs* (@dog_rates).



If they don't, kindly tell them that should follow that handle ASAP because, as you know, they are missing out on daily doses of user-submitted photos and videos of dogs showcasing their adorableness, compassion, courage, and heroism.



Each dog (or "doggo" or "pupper" depending on its assumed age) is rated on a scale of 1 to 10. More times than not however, each dog is given a rating above 10 like "12/10" or the rarest "15/10". The clever young man who runs the handle often precedes each rating with a humorous comment and a fitting dog pun. The handle has been active since 2015 and has gained a massive following on Twitter, Instagram, and Facebook where the posts get thousands of likes and retweets.

I had the chance to analyze some twitter data for an assignment for a data analyst certificate and I'd like to share some of my findings! I gathered data for tweets between 2015 and 2017. Among other data, each tweet had the name of the dog, its rating, the name of its breed (generated by a neural network image recognition), and its retweet count. These metrics were interesting to me so I sought to answer the following questions:

1. What is the most retweeted tweet?
2. What is the most common rating?
3. What are the most common breeds found by the neural network?
4. What is the average retweet count for each rating?

1. What is the most retweeted tweet?

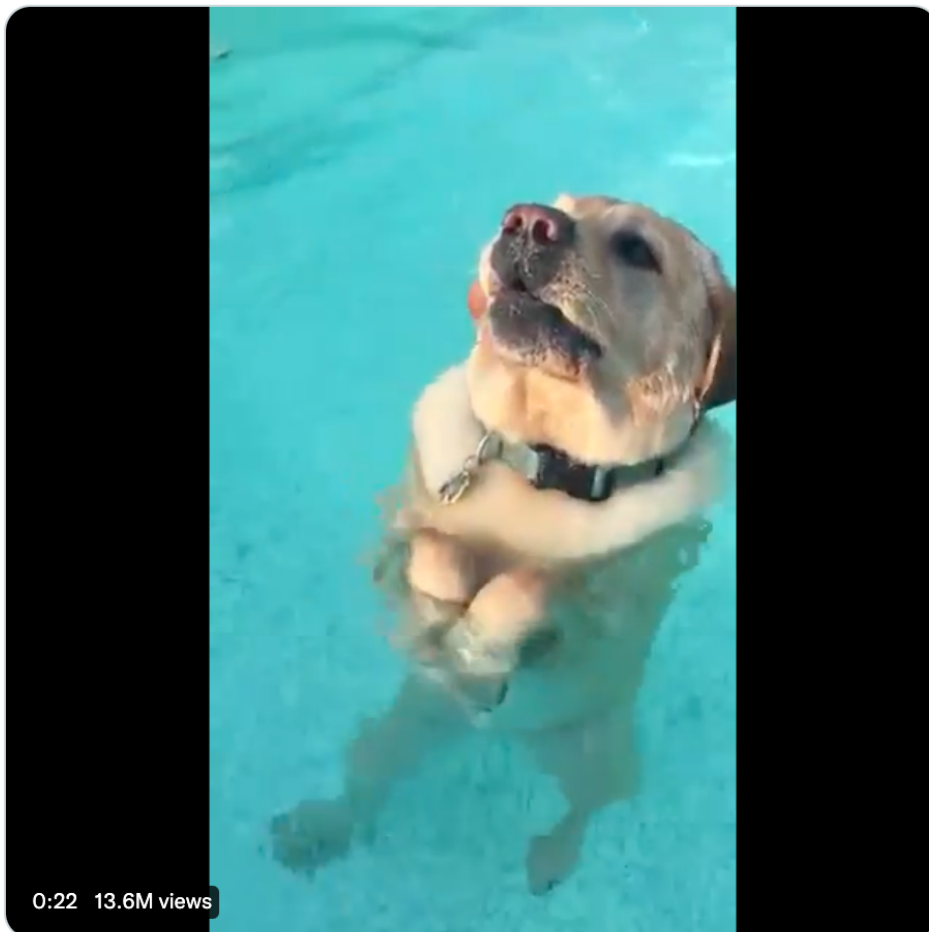
Behold! The most tweeted *WeRateDogs* tweet (between 2015 and 2017):



WeRateDogs® ✓
@dog_rates



Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad)



3:26 AM · Jun 19, 2016 · Twitter for iPhone

72.3K Retweets **3,704** Quote Tweets **147.5K** Likes

This very good boy was doggy paddling in the pool then set his back legs down on the

bottom of the pool and realized he would just stand and not sink. It must've been a surreal realization for the pup, in addition to the sensation of standing without anyone supporting his arms.

If you haven't yet, take a look the original tweet and video [here](#)!

2. What is the most common rating?

Out of the 2000 tweets I examined, surprisingly the most common rating was 12/10! Not 8 or 9, but 12! In fact 70% of the tweets examined were rated 10 or higher.

What's actually surprising is that 30% of tweets had a rating UNDER 10. I've followed *WeRateDogs* for several years and I've never seen a tweet under 12. If I had more recent data, I would wager that the trend was that the average ratings are getting higher and higher. I believe only the early years, before the handle found its rhythm, had actual low ratings. I believe 12, 13, 14, and ONLY in rare cases does a dog get 15 nowadays.

3. What are the most common breeds found by the neural network?

The neural network did well in correctly identifying the breed of dog. Here are the top 5 common breeds by number of tweets.

5. Pug



4. Chihuahua



3. Pembroke (Welsh Corgi)



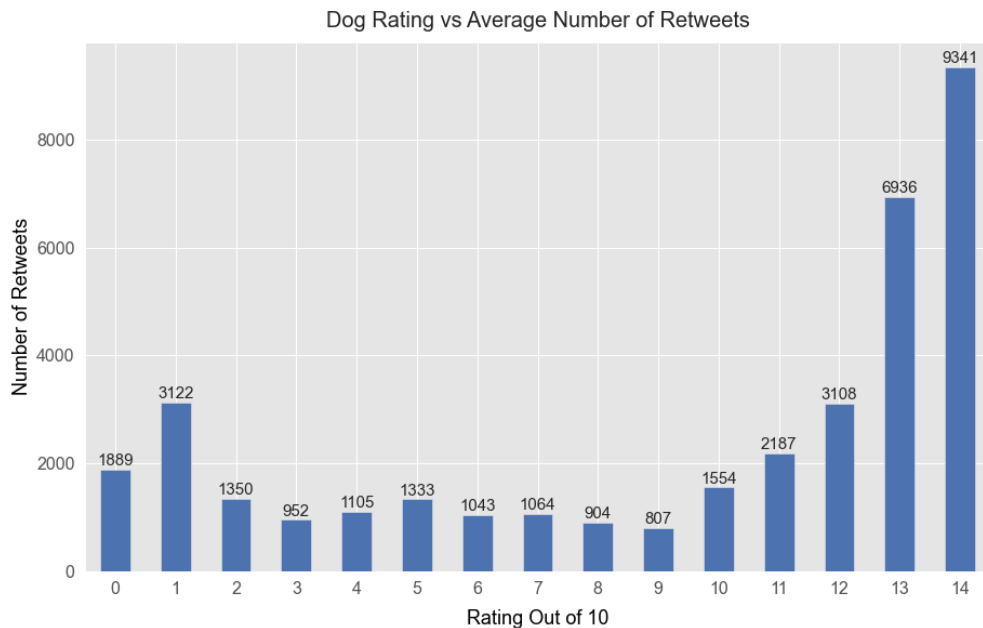
2. Labrador Retriever



1. Golden Retriever



4. What is the average retweet count for each rating?



Final Words

For the visualization, I tried to identify correlations or longitudinal trends but ultimately the scope of data was so limited that the visualizations were inconclusive and lacklustre. I *expected* a high correlation between retweet count and rating, but got an Pearson correlation coefficient r -value only 0.3. Conversely, I tried to show an expected trend in rating verses time yet saw only a slight rise in rating as time went on. So I opted to go with a simpler "Dog Rating vs Average Number of Retweets".

I found that this project had an interesting subject matter, but was limited by the time frame of the tweets. Talking about "most retweeted tweet" between 2015 and 2017 is not a huge feat nowadays as this record probably gets broken on a monthly basis in 2021. Trying to come up with "compelling" and "interesting" insights and visualizations when all I could think up were simple metrics to measure was difficult. This is the first time writing a "blog" style report plus we were not provided examples. The wrangling process itself was very time-consuming. And I'm somewhat of a perfectionist, so I couldn't let certain issues go unaddressed. But I suppose that was good practice and good experience. I hope to get great constructive criticism and feedback so that I can make this report good enough to upload to my github portfolio.