WeRateDogs Tweet Analysis

# Background

[WeRateDogs](https://twitter.com/dog_rates) is a [Twitter](https://twitter.com/home) account that takes an unconventional approach to rating photos, and videos of dogs. He uses a normal fraction system, but the numerator is almost always greater than the denominator. He also uses [nonce words](https://en.wiktionary.org/wiki/Appendix:Glossary#nonce_word) such as "doggo" and "pupper" to describe them. His has leveraged his fame to raise funds to help dogs in need through his [15/10 Foundation](https://twitter.com/15outof10). His web site is <https://links.weratedogs.com/>

I have analyzed several thousand tweets from the start of his account and come up with a number of observations, or insights and created some visualizations to help make some of the observations clearer.

# Insight 1: Dog Names

Of 2356 tweets 1390 give the name of the dog, and there are 929 different names. The "most common" names are Charlie and Lucy with 11 uses each, followed by Oliver, Cooper, Penny and Tucker with 10, 10, 9 and 9 uses respectively.

The insight here is that there is a very large diversity in dog names, with no clear favorite names.

# Insight 2: Breeds and Popularity

From those same 2356 original tweets 1666 have had the breed of the dog identified (this was done using machine learning by a course at [Udacity](https://www.udacity.com/). These 1666 identified dogs represent 113 breeds, and there are some clear leaders for breed popularity, at least in photos posted to this twitter feed. I have prepared two pie charts with this data:

A pie chart with different colored circles

Description automatically generated

Here we see only the top ten breeds represented with their percentage for the whole tweet collection. It is clear that the Golden Retriever is the most common, with 9% of the identified images. This represents 156 tweets.

Overall, there is a lot of diversity in the breeds. With the top 10 breeds represent only 30% of the identified dogs.

# Insight 3: Tweet Age vs Tweet Appreciation

For this insight I combined the retweet\_count and favorite\_count values into a single tweet\_love value by adding them together. As both retweeting and favoriting are indications of appreciation.

I expected that the longer a tweet was around, the more appreciation it would get, however the analysis indicates otherwise.

I'll start with a chart showing tweet posting count by week (each bar represents seven days, but the scale along the x axis is in days since the start of the account):

A graph of blue bars

Description automatically generated

This clearly shows that the account started with a large number of tweets per week, then decreased to a lower frequency. If my hypothesis were true, this next chart would show an equally large spike at the beginning of the account:

A graph of blue dots and red line

Description automatically generated

We don't see this at all, in fact the trend line shows something of the opposite. More recent tweets tend to have higher appreciation.

I suspect this is due to his subscriber count having increased, and Twitter being a very "in the moment" platform, where users don't tend to scroll back through past tweets, but instead to scroll through recent tweets. So tweets are mostly seen in their first few days and rarely after that.

# Bonus Insight: Going Viral

It is interesting to look at the top ten tweets, those that went viral.

Five of the top ten tweets, half of them, contain a video. Three have two or more images and only two contain a single image.

If a picture is worth a thousand words, two or more are worth even more, and a video, even a short one seems to be worth a million words.

Here are links to the top ten, enjoy:

1. <https://t.co/7wE9LTEXC4>
2. <https://t.co/nTz3FtorBc>
3. <https://t.co/DkBYaCAg2d>
4. <https://t.co/UDT7ZkcTgY>
5. <https://t.co/wk7TfysWHr>
6. <https://t.co/AdPKrI8BZ1>
7. <https://t.co/YcXgHfp1EC>
8. <https://t.co/f1h2Fsalv9>
9. <https://t.co/T1xlgaPujm>
10. <https://t.co/g2nSyGenG9>