

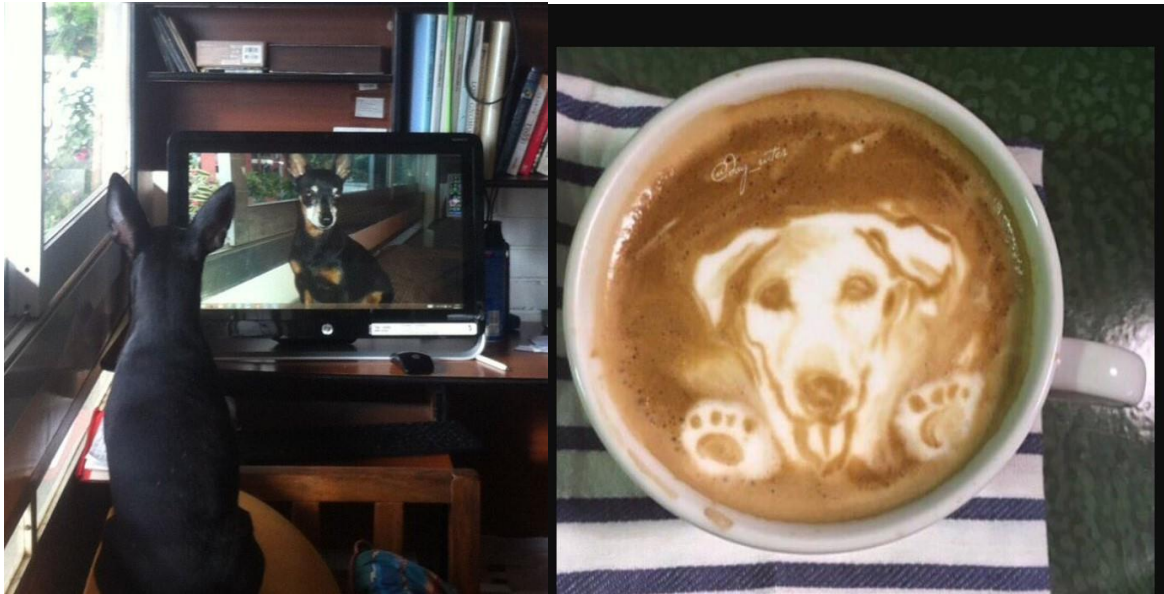
Insight 1: What is the number of puppies compared to doggos and floofers?

First I looked at some basic insights (non-visualized) within the dataset. The most pressing issue to me was the frequency of puppies and doggos, so I called up the value counts of the dog nicknames. Sadly, “None” was the most frequent, but after that came pupper at 234 instances, and doggo at 75. This tracks with accepted scientific convention that most dogs are puppies and the rest are doggos. We also had “puppo” at 25 and “floofer” at 10. I take some academic issue with the inclusion of floofer, as that historically refers to cats. I ran the percentages and found doggo to make up 10% of the total data and pupper to make up 3.4%. It should be noted though that, as reflected in the data, an individual canine can belong to more than one category. If we accept WRD’s questionable application of the term “floofer”, we could accept a floofy pupper being a Samoyed or a Husky.

Insight 2: What is the most common dog breed?

This required utilizing the neural network’s data. Again utilizing value counts we discovered that golden retriever is the most common. That isn’t surprising but it was a bit disappointing because they’re sort of generic (which is not to say that they’re not wonderful, because they are). Also unsurprising was that Labrador came in second place. I was pleased to see a type of Corgi in third, which is a slightly more unconventional choice.

Pulling up value counts also presented me with the bottom several, and I saw some things like “espresso,” “bib,” and “desktop computer”, so I looked for the tweet classified as desktop computer and it was an image of what looks like a Chihuahua mix looking at his image on a computer screen. So the NN made some mistakes, although I suppose that is a reasonable one. The one tagged “espresso” was in fact a cup of espresso with a pretty incredible work of foam art in the shape of a dog.



Insight 3: Who are the highest and lowest rated dogs?

For this I pulled up the highest and lowest numerator values, understanding that even the really high numerators have high denominators, since they're rating multiple dogs. In this case the highest numerator was 1776, which I immediately assumed to be a US-related joke (1776 was when the US declaration of independence was ratified), which it was. This boi received a rating of 1776/10, or a rating of 177.6%. I'm from Panama, not the US, so I don't share that patriotism, but what a cute dog. Incidentally, I feel that he could be considered a floofer based on our species-diverse definition.

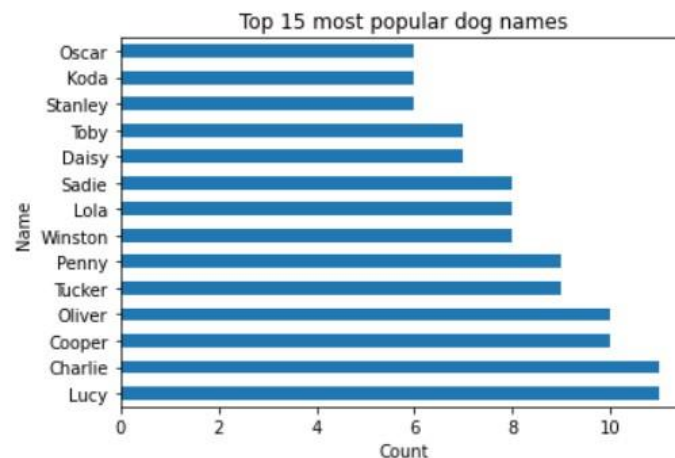


The lowest rated dog had a numerator of 0 which was horrifying, until I pulled the tweet up and saw that it was a tweet chastising someone for plagiarism and the 0 rating actually referred to that person, not the dog in the picture. The dog picture, a golden retriever puppy being pushed in a swing, was intended as a visual pun and was not being rated. I rated him a 40/10 and noticed that the NN categorized him as a swing, not a dog.



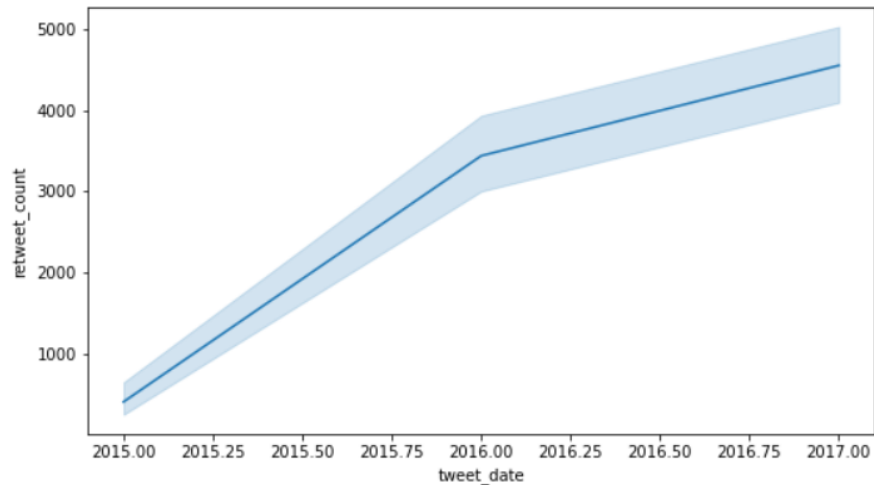
Visualization 1: What's the most popular dog name?

For the first visualization I did a quick bar chart of the most popular dog names. I was again disappointed to find Lucy and Charlie as tied for the most common one (and immediately imagined a golden retriever named Charlie which is so generic). All the names in the top 15 were sort of bland people names. I was relieved to find no Game of Thrones characters in this list.



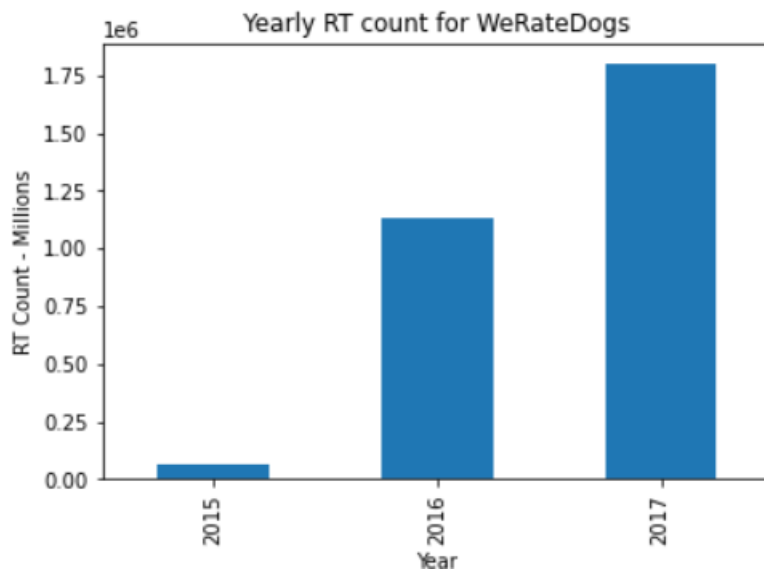
Visualization 2: How has user engagement changed over time?

For this I looked at his retweet count and favorite count. I did a single line chart for retweets:



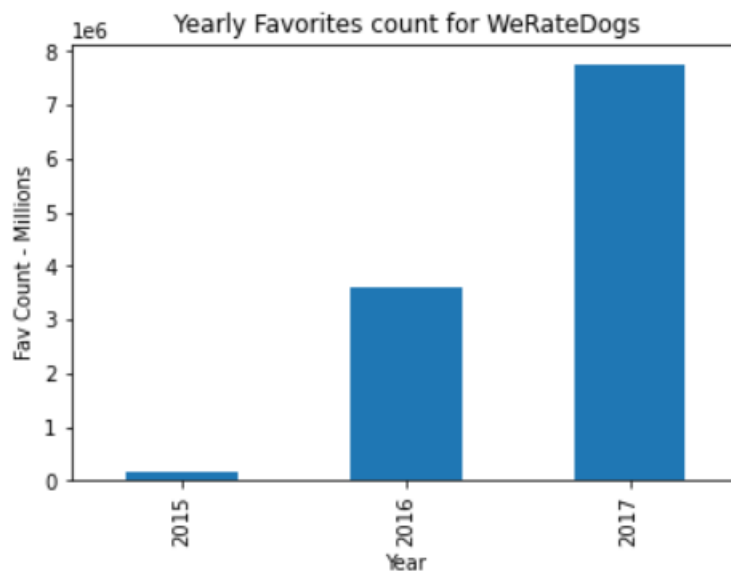
It shows a steep increase up to around 2016 and then a slight slowing in rate.

Then I broke it down into a vertical bar chart showing totals per year (2015, 2016, 2017) which showed almost nothing in 2015, followed by a significant jump in 2016 and a slightly smaller but still quite significant jump in 2017. Essentially confirming the line chart.



Then I looked at his favorites count (the amount of times someone likes a tweet). That showed a similar pattern, although actually more pronounced. That isn't surprising, as more people will like

something than retweet it generally, and by 2017 he was so ubiquitous on Twitter there would be little point in RTing everything he tweets.



In 2017 he got around 30 million likes, as opposed to actually just under 650,000 in 2015. That's really an astonishing rate of growth but then again it's literally one of the best accounts on Twitter.