Act report

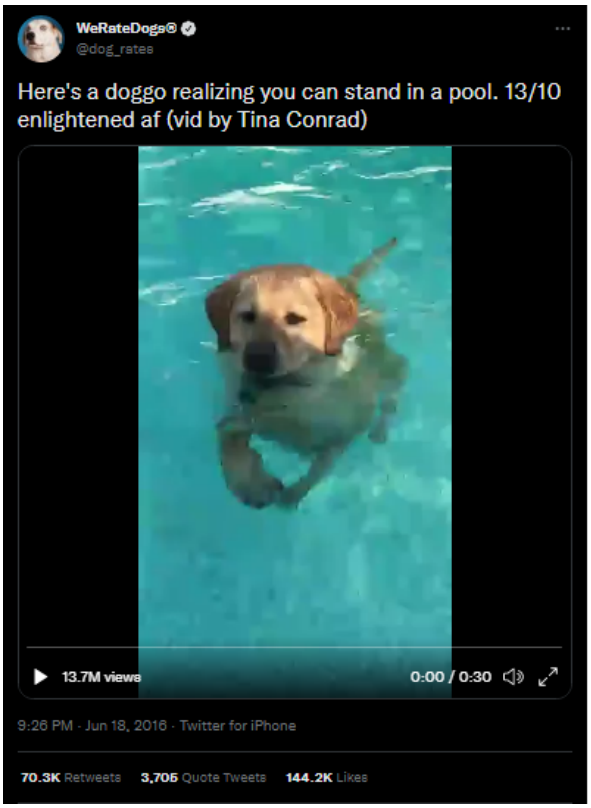
The dataset is from the tweet archive of Twitter user @dog_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dogs. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 13/10, etc. Why? Because "they're good dogs Brent." source

Some of the data were sourced from the WeRateDogs Twitter archive made available to Udacity, where we gathered it by downloading it. Udacity ran it through a neural network to detect the presence of a dog in every image and predicted the breed of the dog where applicable. A result of this was provided to us and downloaded into our workspace. We further queried Twitter's API to enrich our data further.

In our analysis, we asked:

1. Which was the most popular tweet through favourite count and retweet count?
2. Which was the most common dog breed?
3. Was there a relationship between the rating_numerator and the favorite_count?

We found that the tweet with the tweet_id: 744234799360020481, tweeted on 2016-06-18 at 18:26:18, was the most favourited (144310 times) tweet and retweeted, (70366 times) in our dataset. https://t.co/7wE9LTEXC4



The Golden Retriever was the most common dog breed in our dataset with 156 entries. Finally, we noted that there is a weak positive correlation between the favourite count and the rating numerator at 0.07 and plotted a scatter plot to visualize it.