Twitter archive for an unconventional way to review and rate dogs, had 1,400 records about different breeds of dogs, their names, and the stage(Pupper, Doggo, Puppo, or Floofer)associated with some dogs in the archive. The archive had gone through a series of wrangling steps, where it was cleaned to increase the quality of the content, and add to it more relevant data to potentially extract useful insights.

Such relevant data was the breed of dogs that was extrapolated from dog images associated with each tweet in the archive using an Artificial Neural Network(ANN). Using this data, it was possible to create the visual displayed below which showcases the distribution of the different dog breeds throughout the tweets history. During analysis of the tweets history, it was discovered that there were 111 different dog breeds associated with 1,400 dog review tweets. The visual below, however, shows the top ten most tweeted about dog breeds in that particular archive. And the first spot seems to go to the Golden Retriever breed, followed by Labrador Retriever, and then Pembroke.

Another source gathered for more relevant information about the archive is the engagement metrics associated with each tweet, such as, how many people have liked or re-tweeted each tweet. And during the analysis, it was discovered that "pupper" and "doggo" dog stages were associated with tweets with the highest engagement. It's worth mentioning that the source of this supplementary data was through the Twitter API. The API was queried using the Tweet Id associated with each tweet in the archive and retrieved the "favorite_count" & "retweet_count" attributes of a given tweet. As for the data about the dog stages, it was provided in the same archive but had to be restructured to be able to perform the previous analysis.

Alongside the dog breeds inferred by the algorithm, it also provides a measure between zero and one for how much confident the algorithm is, with its prediction(the closer the measure is to one, the more confident the algorithm would be in its prediction of the dog breed from the image). That, along with the engagement metrics, it was possible during the analysis to identify the tweets with the highest engagement. As well as the prediction accuracy, on average, associated with these tweets. And the average prediction accuracy measured for these tweets was around 0.65.

It was also interesting to know, what dog breeds were the most accurately predicted by the algorithm, and during the examining of this question it was found that the following dog breeds had the highest chance of being represented by the algorithm:

1. komondor

2. clumber

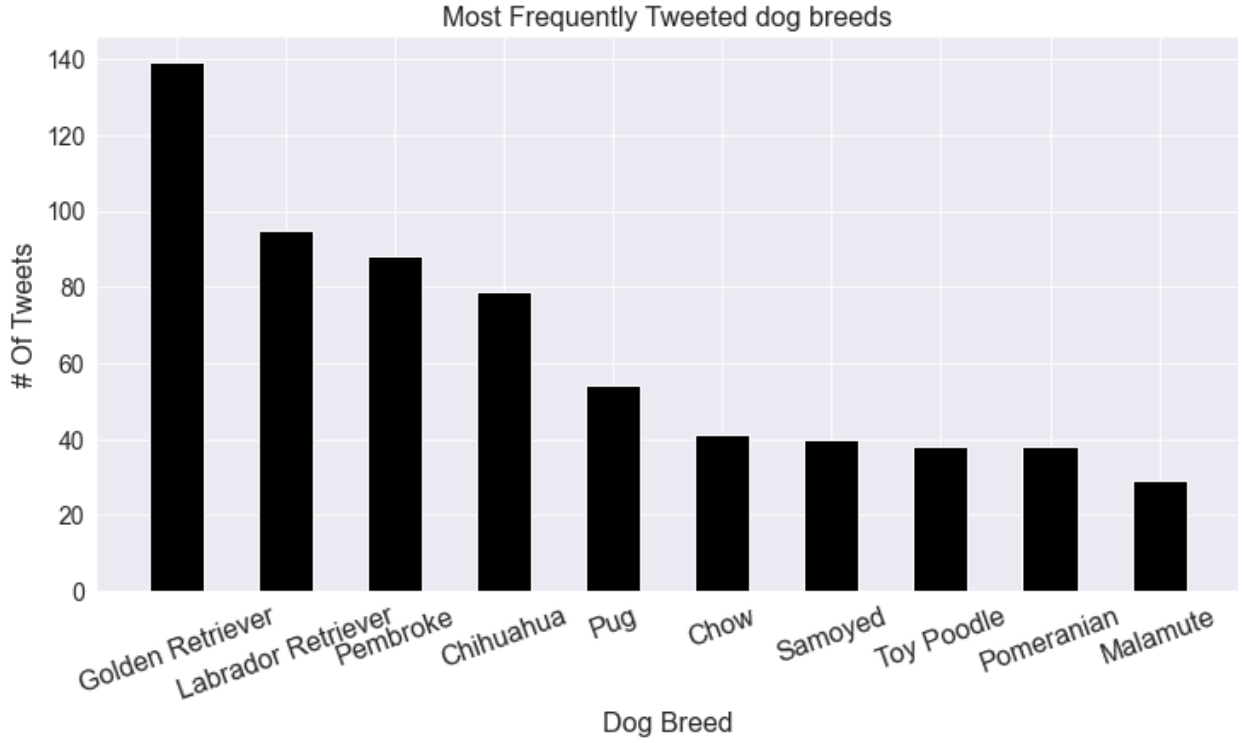
3. Brittany spaniel

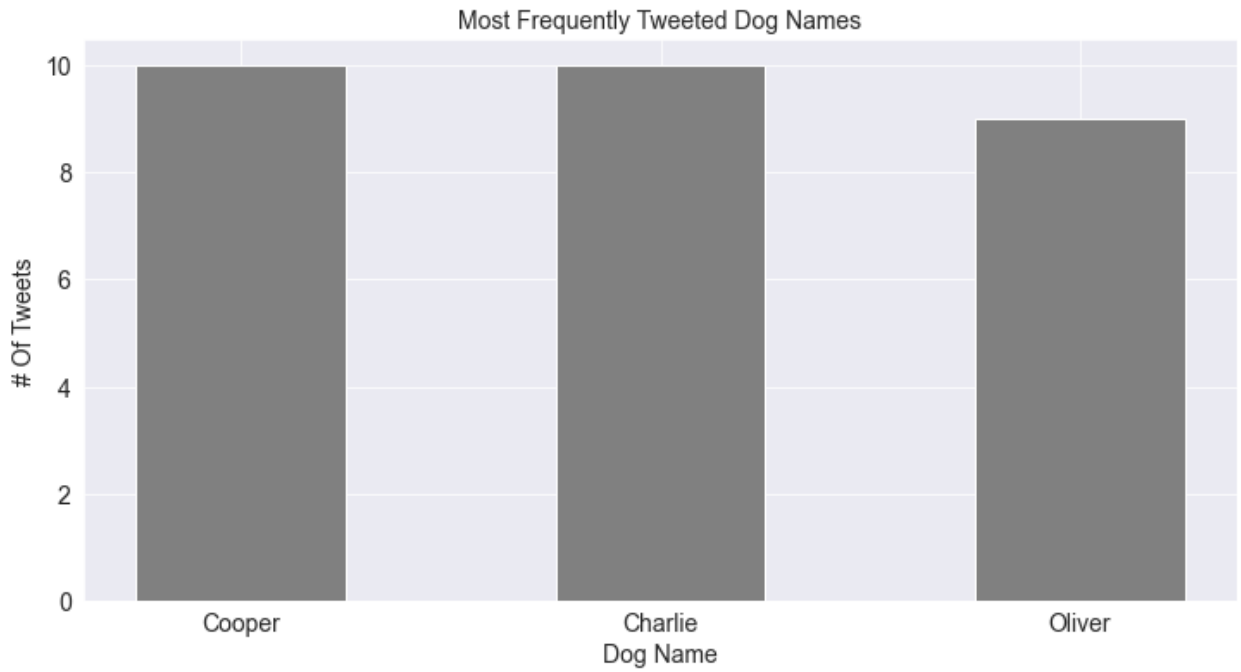
4. Keeshond

5. bullmastiff

6. Bernese mountain dog

Based solely on the archive dataframe and despite the presence of many names for dogs as opposed to the dog breeds. An analysis for the most common dog names showcased the second visual plot in this report where it displays the three most common names for dogs in that particular archive that spans two years worth of dog reviews, and as shown, the first most popular dog name is "Charlie" followed by "Cooper" and then "Oliver."




Most Frequently Tweeted Dog Names
10
8
6
4
2
0
Of Tweets
Cooper
Charlie
Oliver
Dog Name