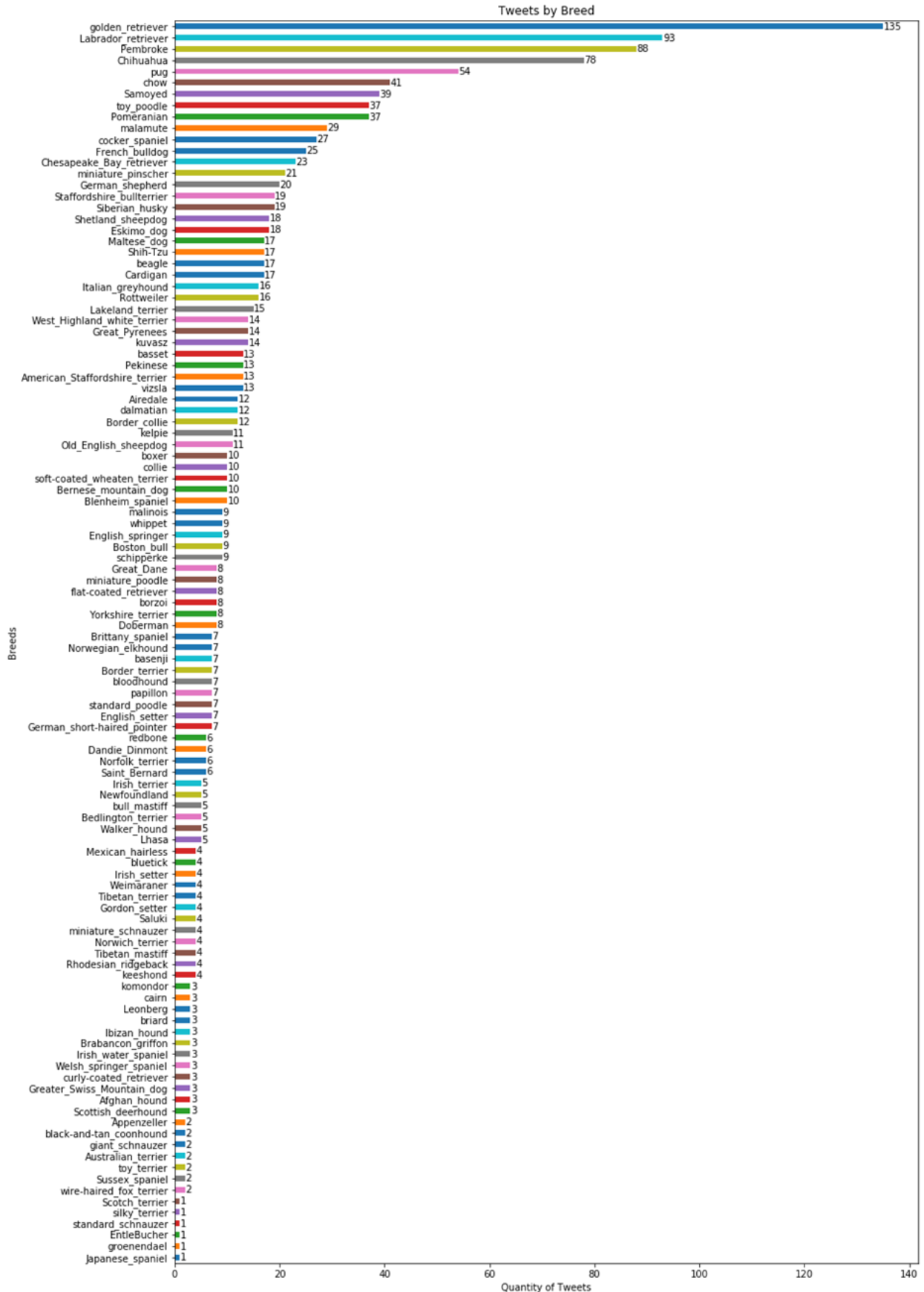


Exploring WeRateDogs

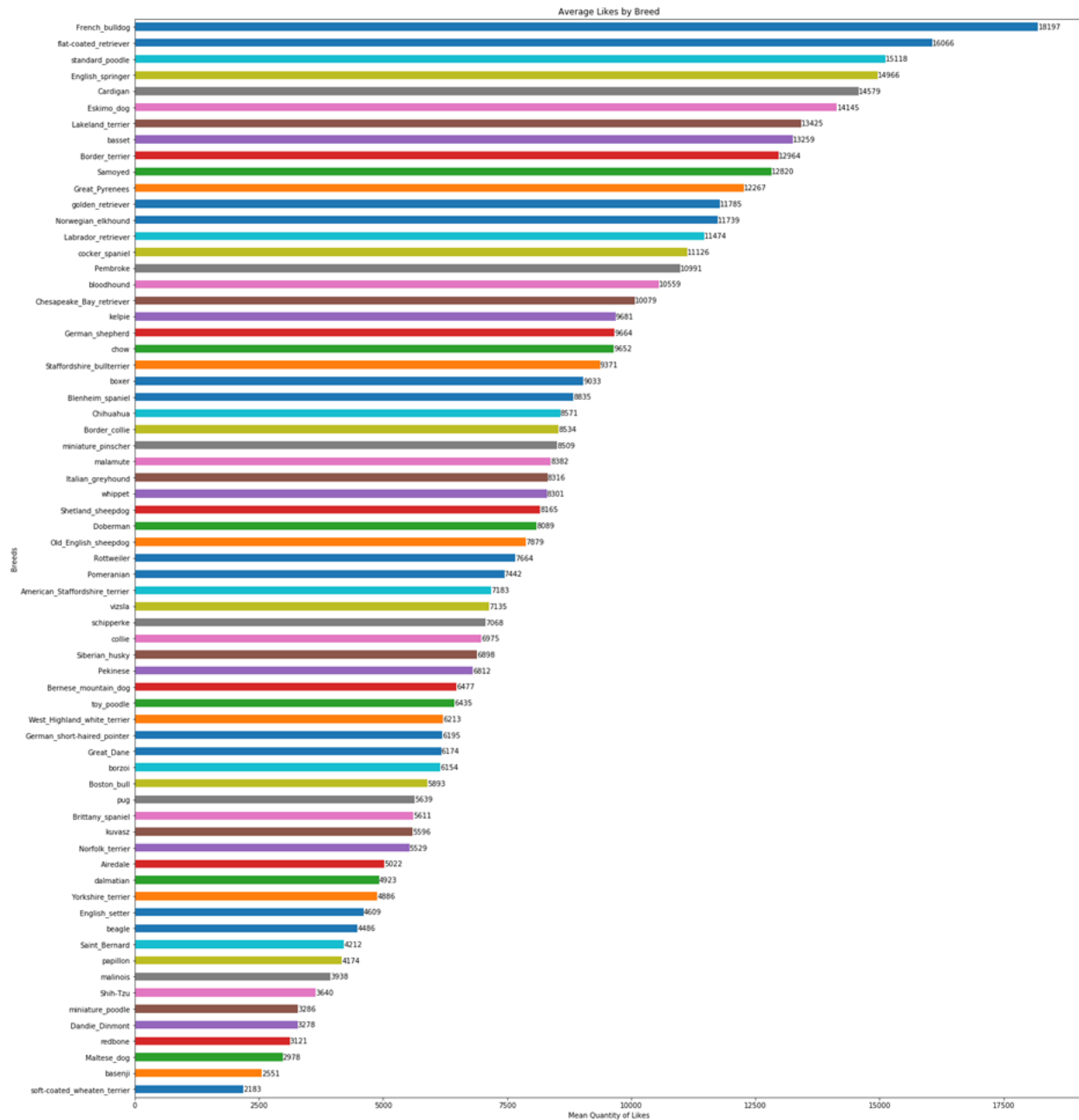
The exploration of the gathered WeRateDogs data mainly involved an analysis of the various breeds, including charting the number of tweets, likes, retweets, and ratings by breeds; as well as an analysis of correlations in dataset, including positive correlations with the ratings score.

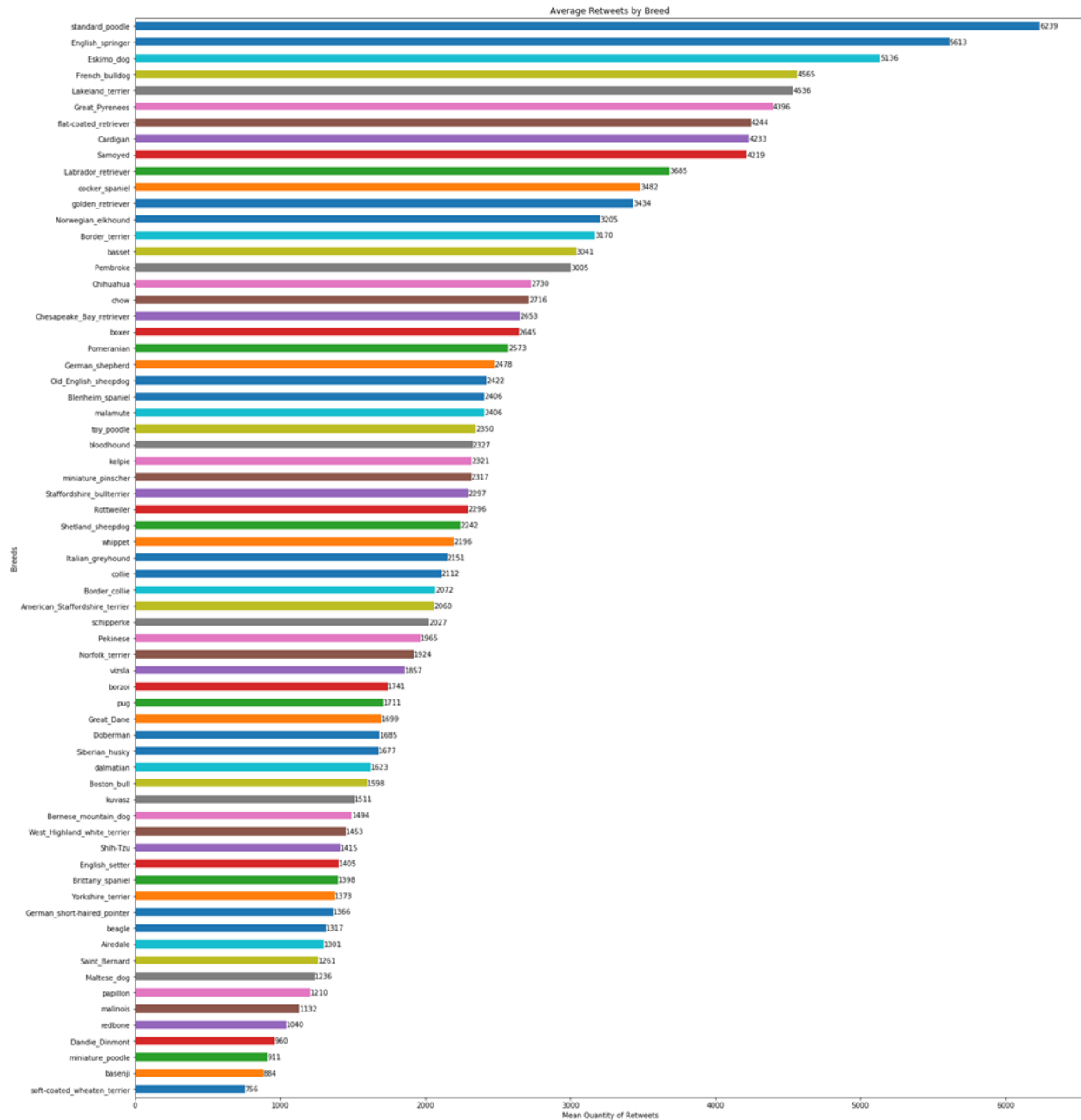
Breed Analysis

To begin, I plotted the quantity of tweets for each breed so I can which breeds were popular. As seen below, perhaps unsurprisingly, Golden Retrievers and Labrador Retrievers were the most popular. I'm guessing that the number of tweets just reflect how common the breeds are in general.

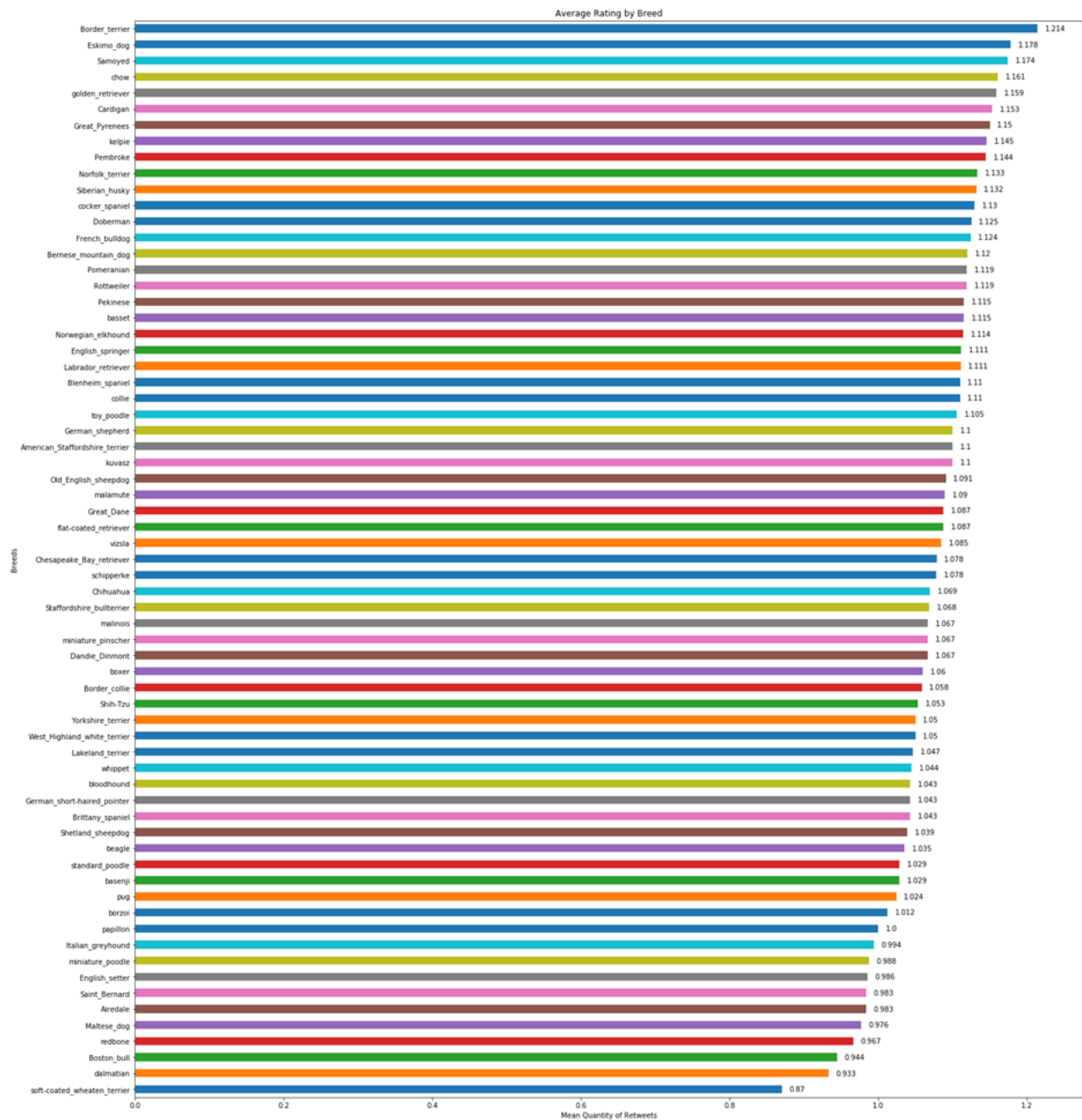


While the last result was unsurprising, perhaps there are other breeds that are more popular likes and retweets-wise. So, for breeds with more than 5 tweets, I plotted the average likes and retweets for each breed. Here, we see that Golden Retrievers are not at the top. French Bulldogs generate the most likes, and Standard Poodles generate the most retweets.





Just because the data was available, I also plotted the average ratings. Apparently Border Terriers are the best dogs among very good dogs.

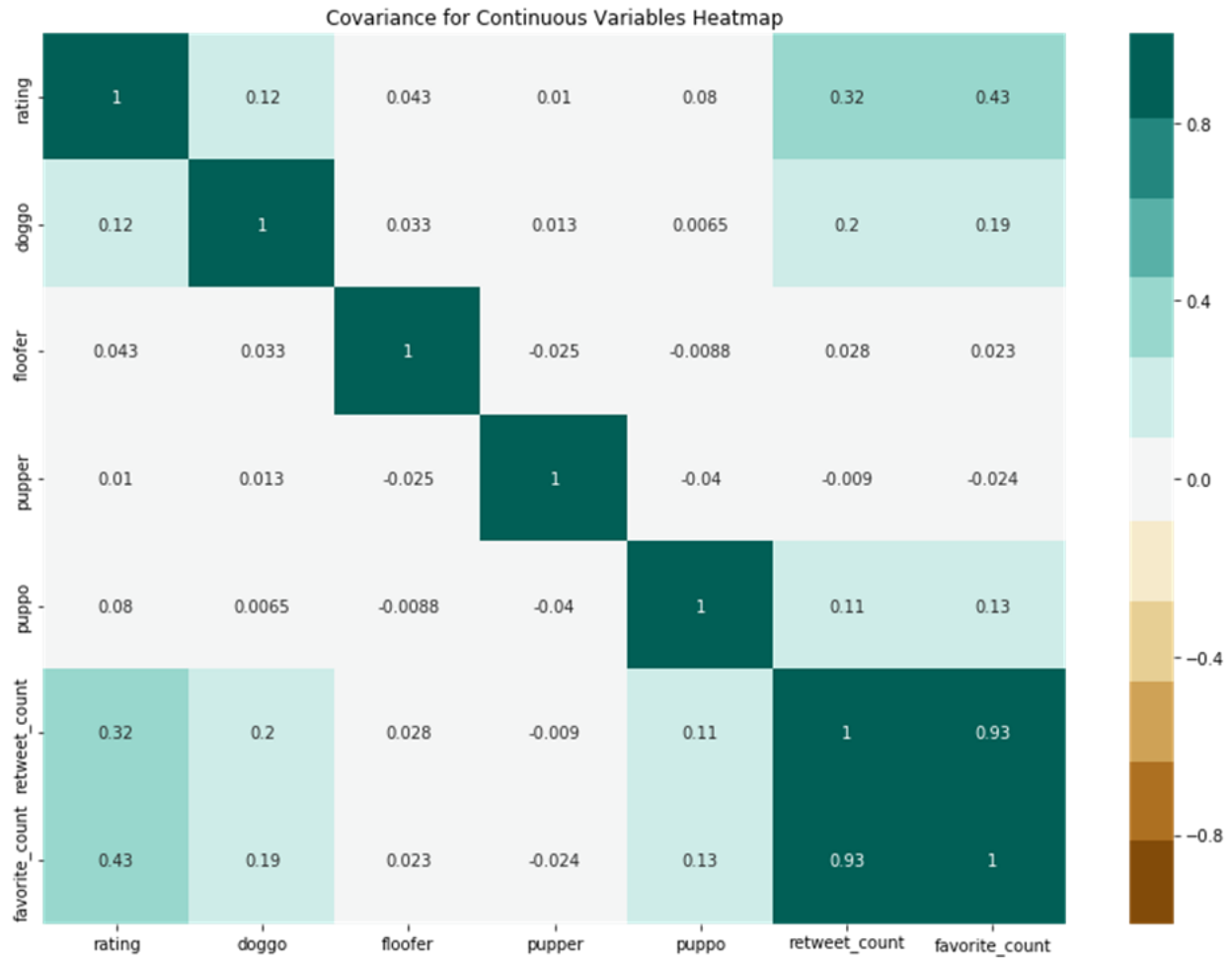


Exploring Correlations

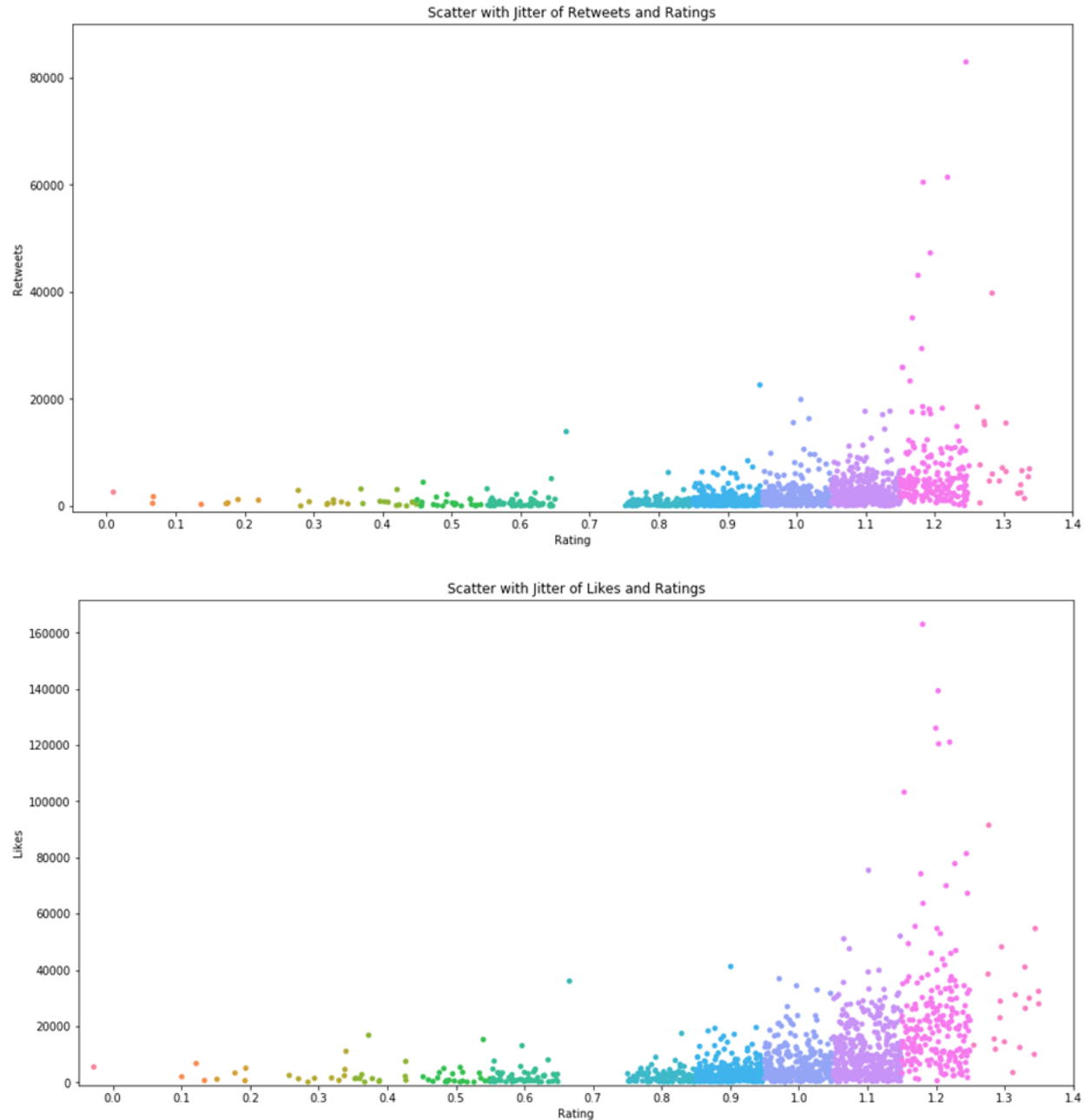
Given there were a number of continuous variables in the dataset, I generated a correlation-table heatmap to calculate and visualize significant correlations.

From reviewing this table, I found that (1) of the dog “terms of endearment” (“puppo”, “pupper”, “floofer”, and “doggo”), only “doggo” and “puppo” generated notable correlations, and that (2) ratings appear to have a significant correlation with likes and retweets.

For (1), “doggo” and “puppo” appear to slight positive correlations with likes and retweets. I guess the insight here would be that it wouldn’t hurt to tweet “doggo” and “puppo” more often (or post more pictures of “doggos” and “puppos”? I’m not sure what these words mean).



For (2), given that significant correlations were found for the ratings, I wanted to visualize the relationship to see if the correlations were visible. Even though the ratings were technically continuous values, they were mostly discrete and falling mostly about 0.1 increments. Therefore, I scattered the variables with jitter so the points didn't overlap and misrepresent the quantity of points at or near a particular coordinate. Funny enough, it does appear that the higher ratings are associated with more retweets and likes.



Findings

To recap, the exploration produced the following insights:

1. Although Golden Retrievers are by far the most common breed on WeRateDogs, there are other breeds that garner more likes and retweets.
2. Whether a tweet has a "floofer" and "pupper" doesn't seem impact likes and retweets, but there is a slight positive correlation for "doggos" and "puppos".
3. Higher ratings do appear to garner more likes and retweets.

So maybe for the next rating, WeRateDogs should seek a very good French Bulldog and award the "doggo" a 14/10.