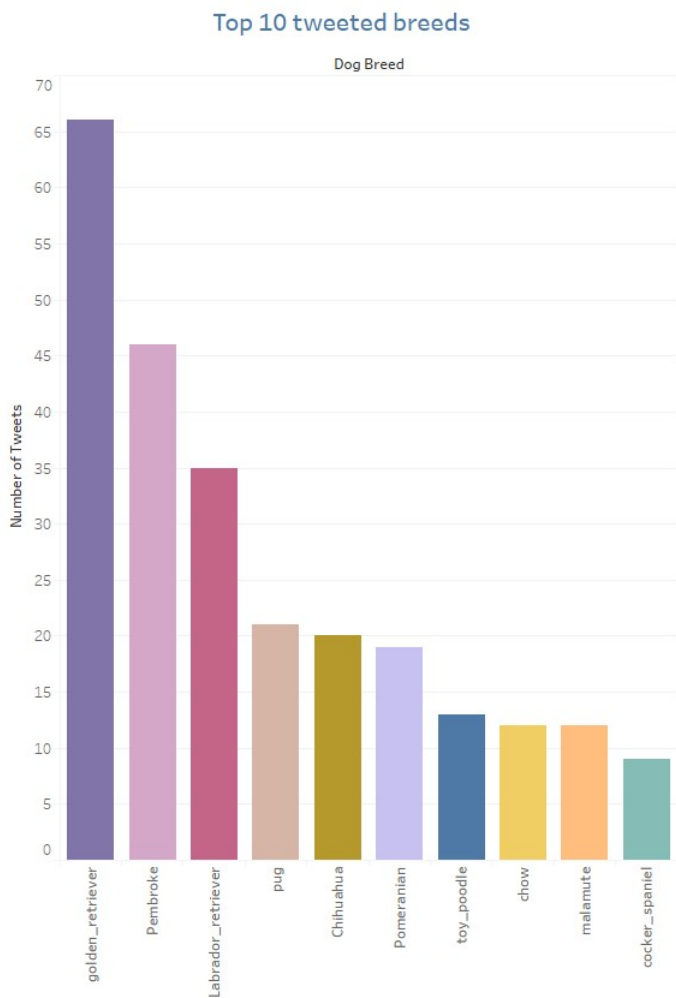


Observations on Twitter Archive data

By Jennifer Powell

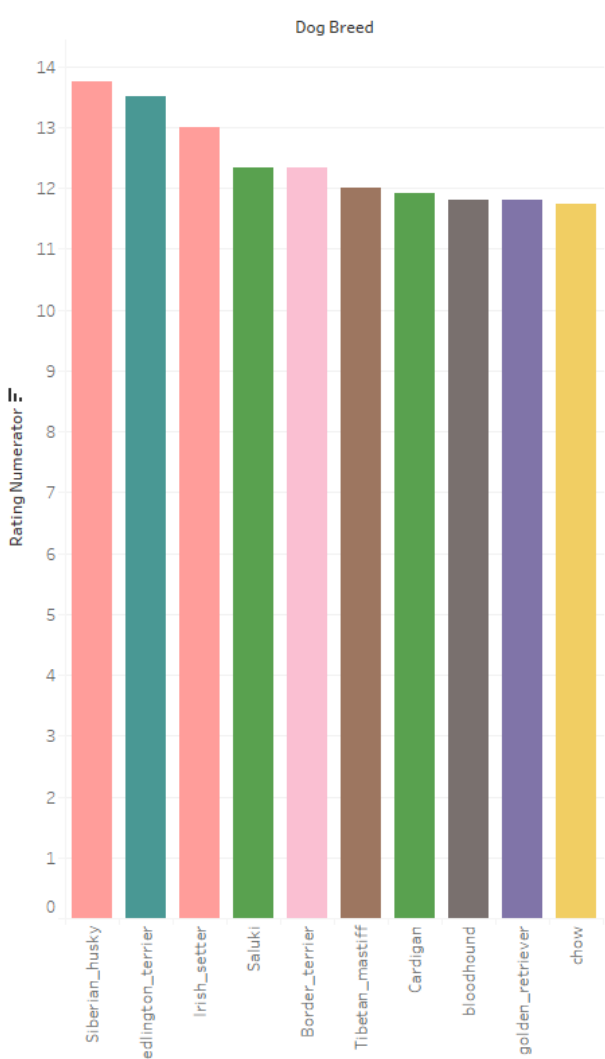
After analyzing the cleaned and wrangled data, I came up with many observations about the data. I will convey the 5 that I found most interesting.

1. The top 10 most tweeted breeds, ie taking a count of all tweet records by breed (P1), and filtering for p1 confidence greater than .50, I came up with the following list. The Golden Retriever outpaces the next breed (Pembroke) by almost 20 entries.

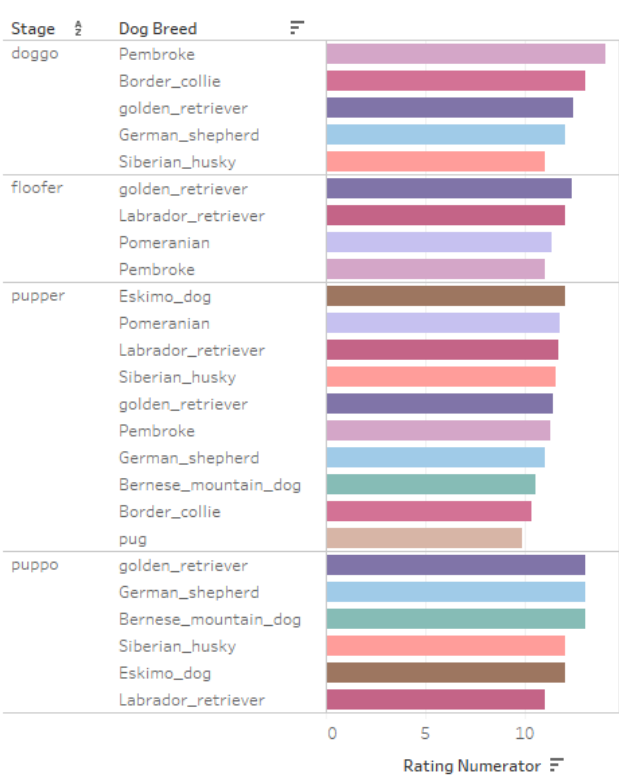


2. When looking at the numerator rating (i.e. x out of 10), the Siberian Husky is at the top with an avg numerator of 13.75 out of 10. When breaking it out by highest numerator per stage, then the Golden Retriever wins 2 out of 4 (flooper and pupper). The Pembroke and Eskimo Dog win for doggo and pupper.

Top 10 Avg Numerator Rating by breed

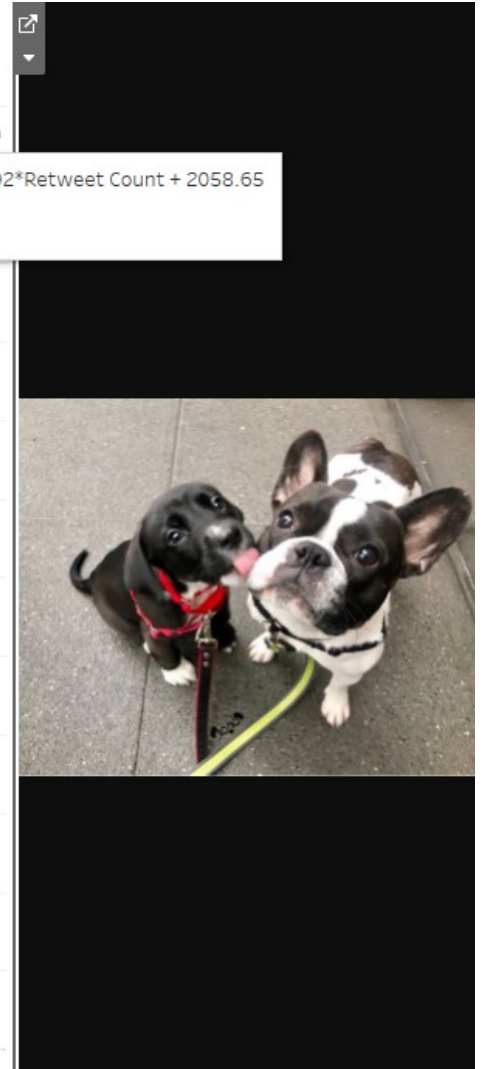
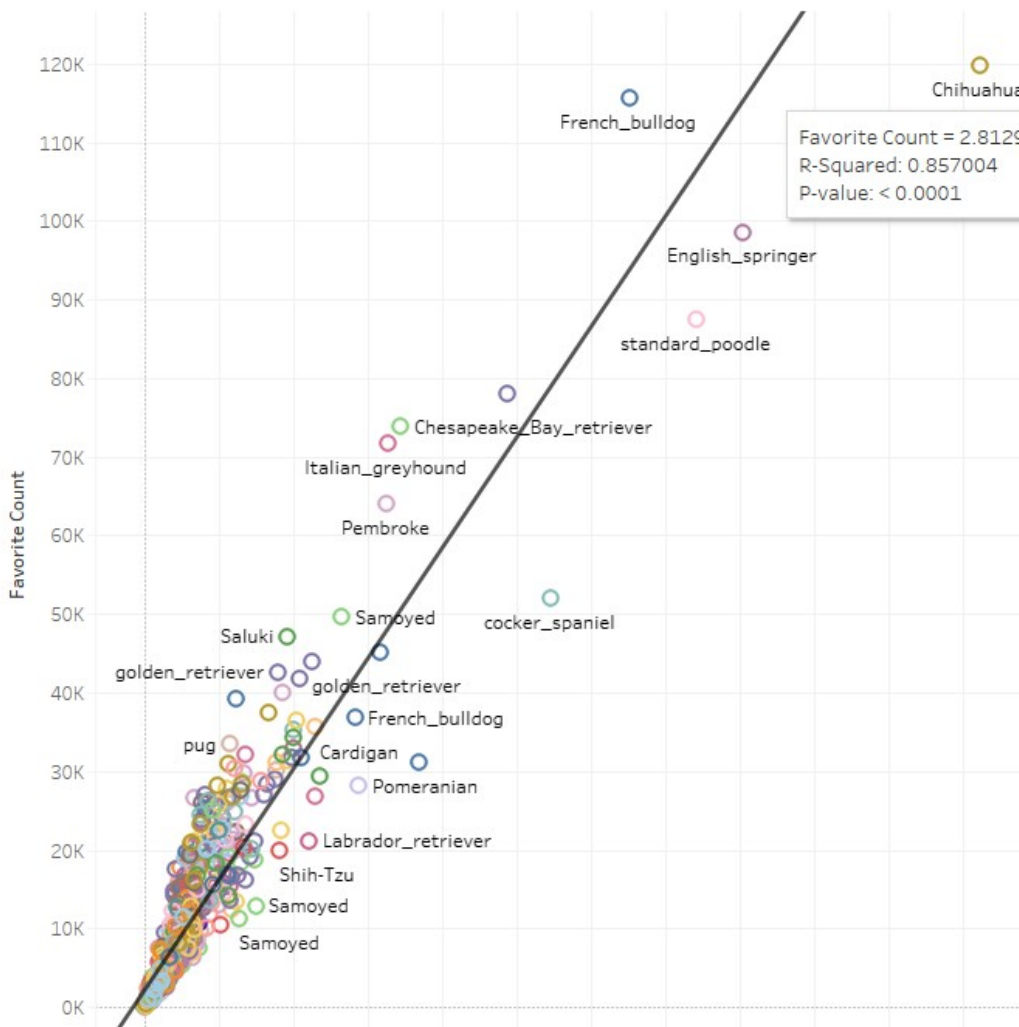


Avg Numerator rating by breed per stage

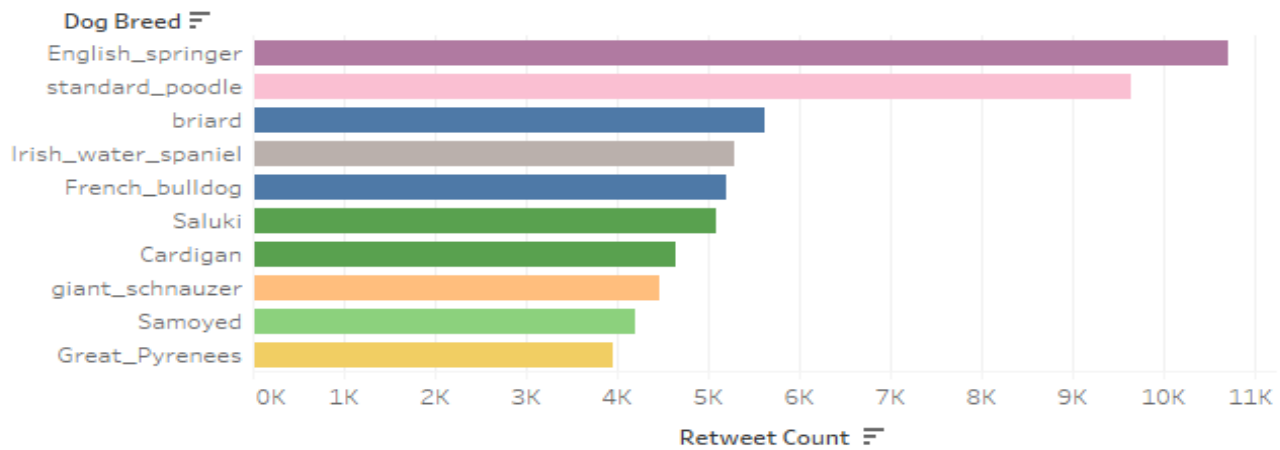


3. In looking at the Favorite Count and the Retweet Count, there is a definite correlation between the two. The r-squared value is .85 and the p-value is significant at $< .0001$. If you look at the separate charts (top 10 avg for each) there is a lot of overlap. English Springer Spaniel, and Standard Poodle and Irish Water Spaniel are all at the top on both charts. So, you could say with some certainty, that a dog breed with a high favorite count, will most likely also have a high retweet count.

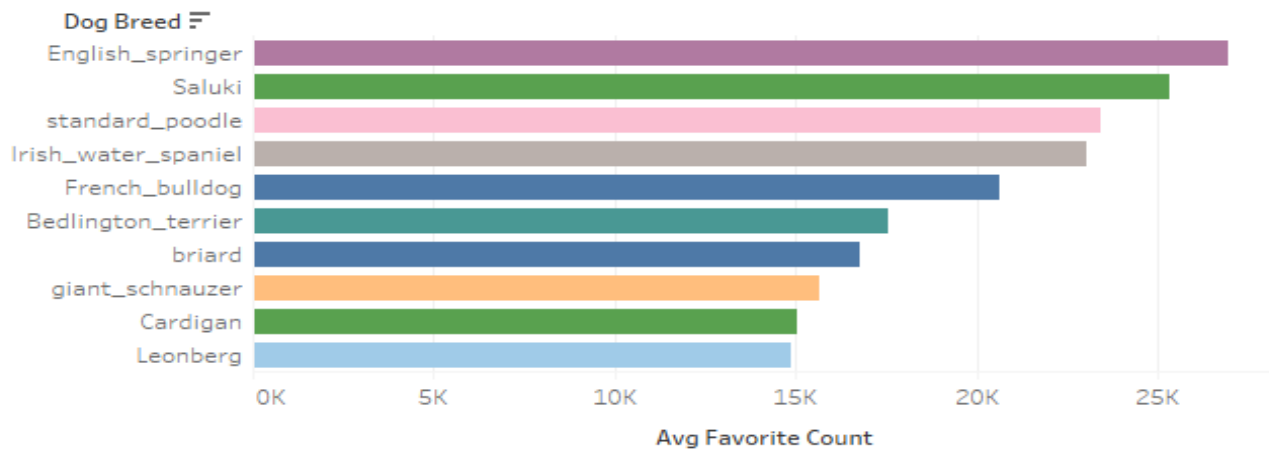
Favorite Cnt vs Retweet Cnt



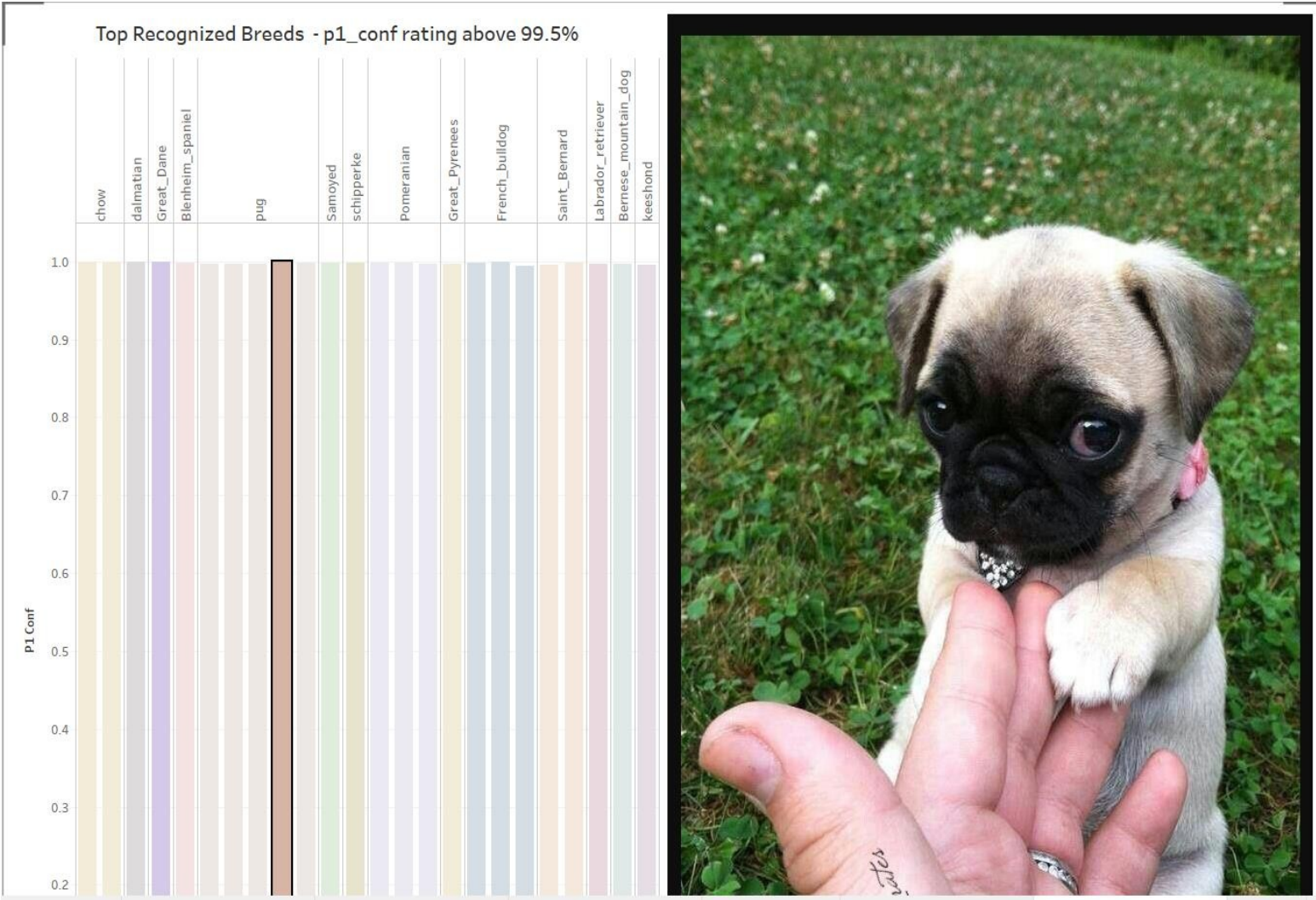
Top 10 Breeds by Avg Retweet Cnt



Top 10 Breeds by Avg Favorite Cnt



4. Another interesting thing I found, was the accuracy rate of the machine learning that tried to determine the breed. Since I'm using dog breed as a parameter for many of these observations, I thought it would be interesting to see which breeds were most accurately depicted. In clicking through the links below they looked to be correct.



5. Last but not least, in looking at dog names, here are the top 10:
 - a. Cooper, Charlie, Oliver, Koda, Sadie, Bo, Leo, Toby, Tucker and Winston

