

Analysis Report

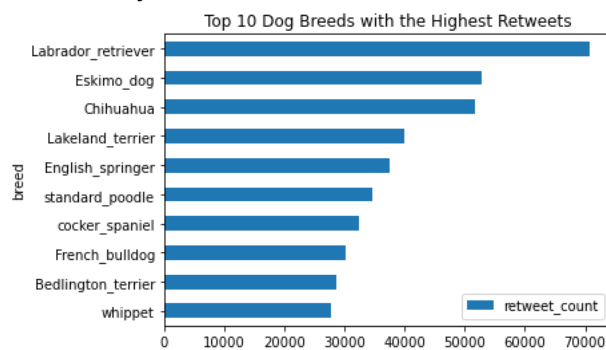
After gathering, assessing and cleaning the datasets, they were merged into one dataset and stored. The goal here was to then analyze the merged data to visualize and get some insights.

Questions which I wanted to answer were:

- Which dog breeds have the highest number of retweet count?
- Which dog breeds have the highest number of favorite count?
- Which breeds are top rated?
- How have tweets fared over time
- Which names of dogs are the most popular
- Which dog types are the most popular
- Which dog types have the highest rating
- How do retweets and favorite relate?

Question 1:

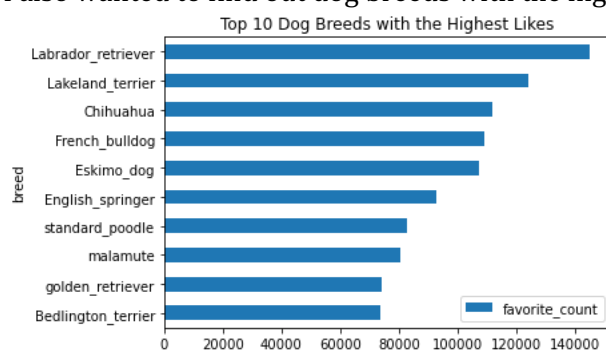
From the created dog_breed column, I wanted to find out the dog breeds that had the highest retweets by other Twitter users.



From the image above, it is evident that the Labrador Retriever was the breed with the highest number of retweets.

Question 2:

I also wanted to find out dog breeds with the highest number of likes by other twitter users.



Here, Labrador Retriever also had the highest number of likes. This infers that a good number of twitter users that engage with the WeRateDogs account have some sort of liking of the Labrador Retriever dog images or posts.

Question 3:

I also wanted to find out the dog breeds that got the highest ratings on average. I first segmented this data to represent data whose rating had a denominator that was equal to 10. The dog breed with the highest rating on average based on this segmentation was Clumber.

Question 4:

For this, I wanted to understand how the WeRateDog tweets have fared over the course of time represented in the dataset. Breaking this down my months, saw a decline, then a bit of stability and finally a decline in the last month represented.

Question 5:

The most popular dog names with counts of 10 each are: Cooper, Charlie and Oliver

Question 6:

The most popular dog type is pupper as shown below:

=

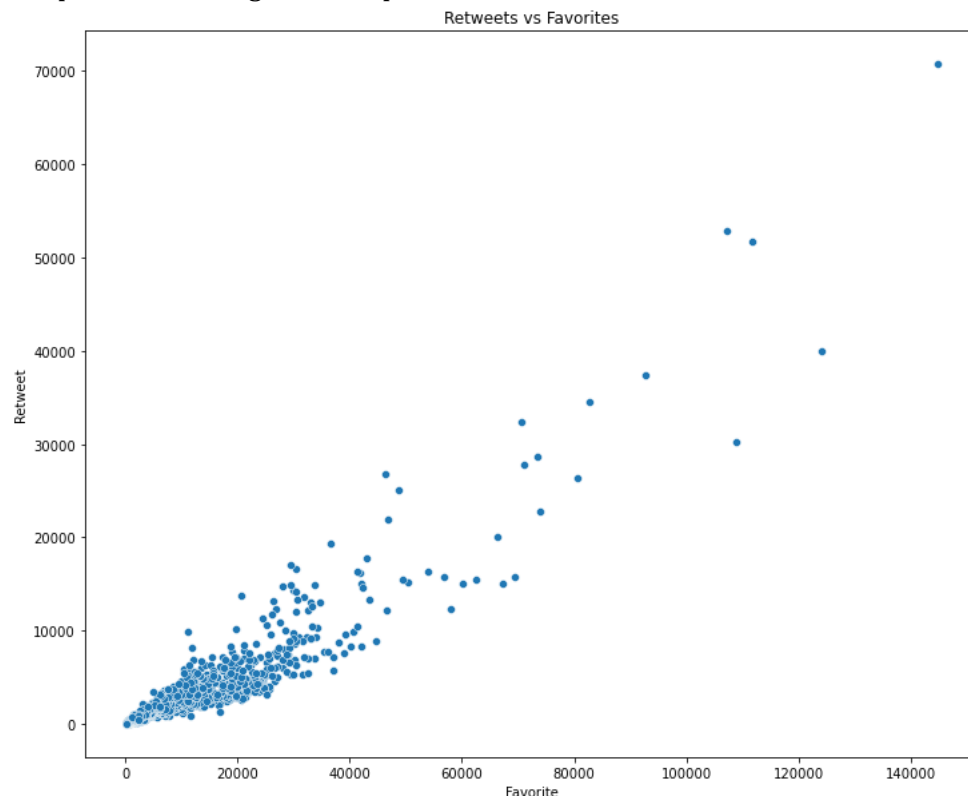
Dog Type	Count
Pupper	222
Doggo	71
Puppo	27
Floofer	3

Question 7:

I also wanted to find out the dog types that got the highest ratings on average. I first segmented this data to represent data whose rating had a denominator that was equal to 10. The dog type with the highest rating on average based on this segmentation was Floofer with a rating of 12.67.

Question 8:

Lastly, I wanted to compare the relationship between favorites and retweets. I did this comparison utilizing a scatter plot.



The plot showed a direct relationship between favorites(likes) and retweets indicating that more likes may result in more retweets.