

Most likeable breed

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. In this project, we will find out which breed is the most liked breed by the public.

About dataset

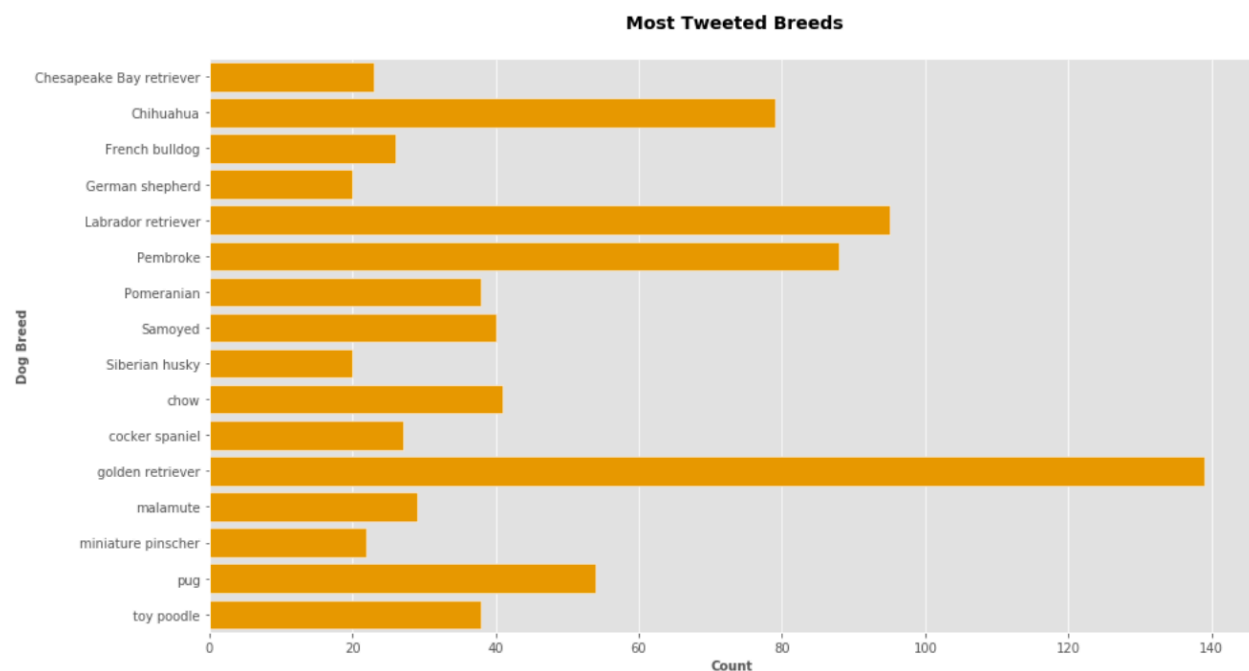
The dataset contains the following files:

1. twitter_archive_enhanced.csv - This file contains 2356 observations of original tweet like - time, source of tweet, comments, urls, different dog stage and name of the dog.
2. image_prediction.tsv - Downloaded from cloud, this file contains 2075 observations of tweet image predictions, i.e., what breed of dog (or other object, animal, etc.) is present in each tweet according to a neural network. This file contains the result of top 3 algorithms. For the purpose of my analysis, I have chosen the 1st prediction algorithm as it had the highest confidence ratio.
3. tweet-json.txt - Extracted from Twitter API, this file contain 2354 observations of favorites and retweet counts of each tweet_id.

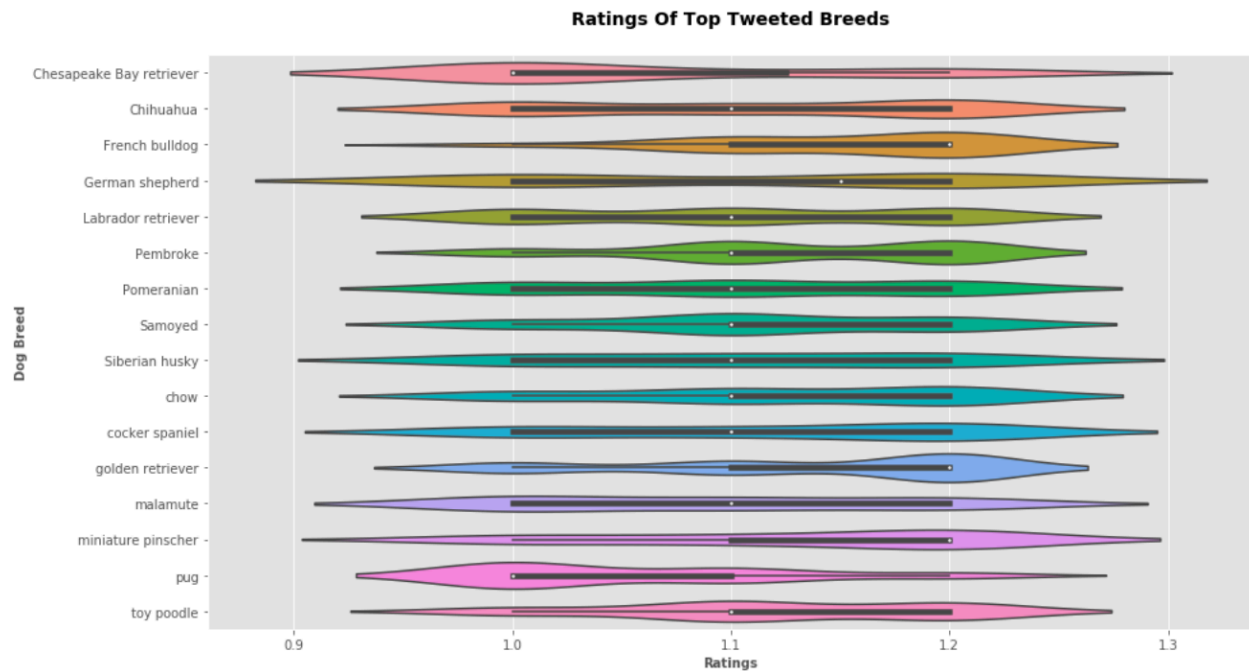
After cleaning the data, we have 1994 observations to analyze.

Which breed is the most likeable breed?

There are many dog breeds. In order to see which breed is the most likeable one, I have filtered the data to choose the breeds that have minimum 20 tweets.

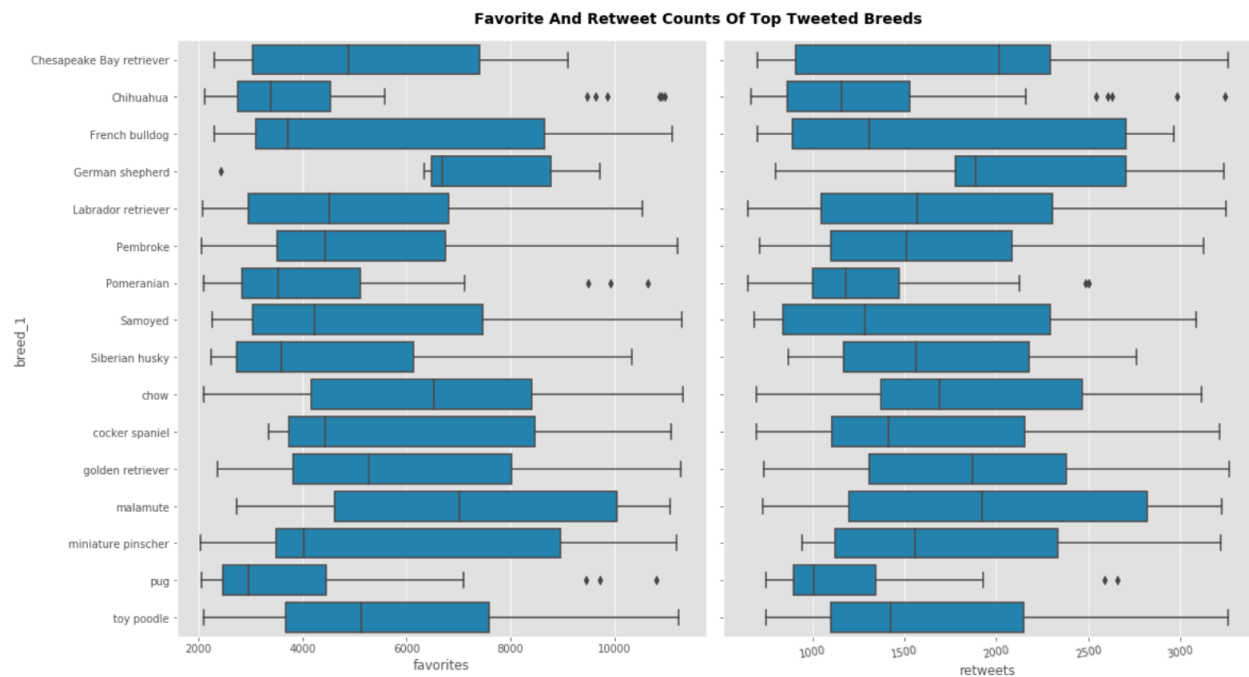


As you can see below, Golden Retriever is the most rated dog breed among others. This plot had me thinking why these breeds are the most tweeted breeds. What makes them so special? I dugged little deeper and plotted the ratings of these breeds.



As you can see, German Shepherd has the highest rating of 1.3 and it's median rating is 1.15. Although Golder Retriever didn't score the highest rating, it's median rating is 1.2 which is more than German Shepherd and it makes Golder Retriever more likeable breed.

While the above plot gave me an idea that Golden Retriever could be the most liked breed, I actually wanted to see if this is true. I plotted number favorites and retweets count for each of the top tweeted breeds.



Hmm! Interestingly, the median favorites count of Golden Retriever is 3000 which is below than German Shepherd who has 6500 favorites count. Again, German Shepherd beats Golden Retriever by 100 retweets. Looks like we know the winner of the most likeable breed!

