

WeRateDogs Twitter Data Insights

by Ting-Hsiang (Lawrence) Tseng

Introduction

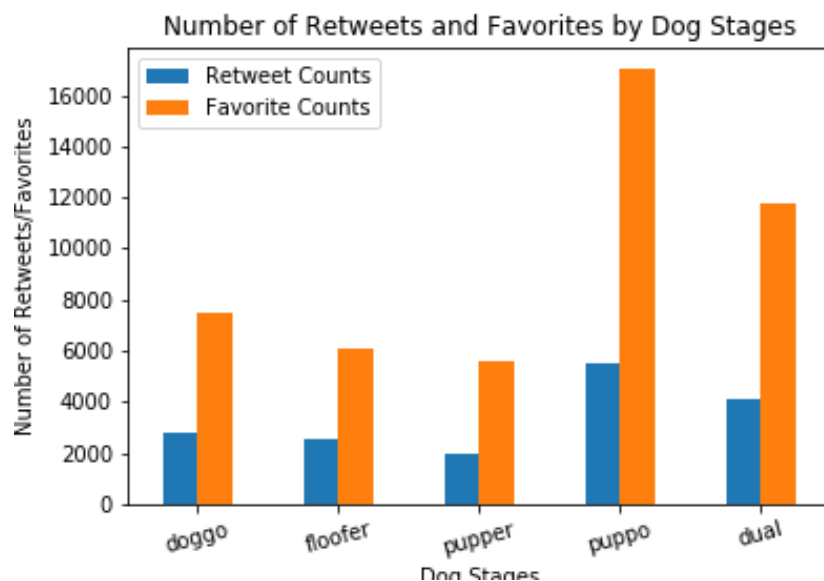
WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. This user has over 4 million followers and received international media coverage.

The dataset that we will be analyzing is *twitter_archive_master.csv*, which was retrieved from three datasets, the WeRateDogs Twitter archive, the tweet image predictions, and retweet/favorite count data obtained by using Twitter API. The dataset of *twitter_archive_master.csv* has been assessed and cleaned. The information within the dataset includes tweet ID, source, content text, dog stage, rating score, tweet's image, breed prediction, and retweet/favorite counts.

Some insights that we found from the *twitter_archive_master* dataset are shown as follows.

1. Popularity of Different Dog Stages

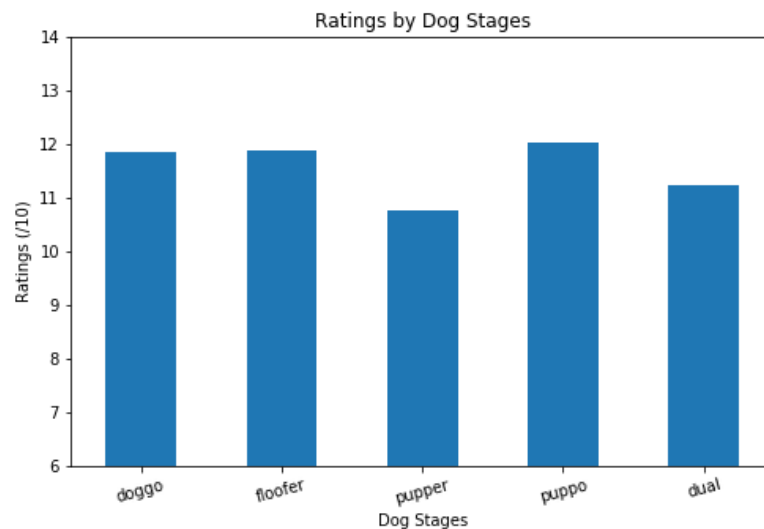
Some tweets in WeRateDogs contain dog stage, and there are four major stages with a few dual (might be incorrectly entered) stages. We are going to see which dog stage has the highest number of retweets and favorites.



According to the graph, we know that the tweets of puppy usually have more retweets/favorites, and the tweets with doggo lingo have the lowest number of retweets/favorites in average.

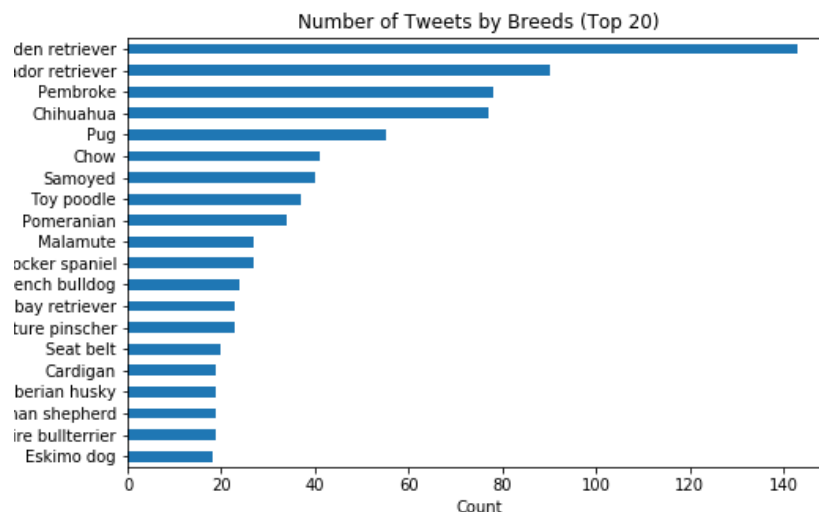
2. Ratings of Different Dog Stages

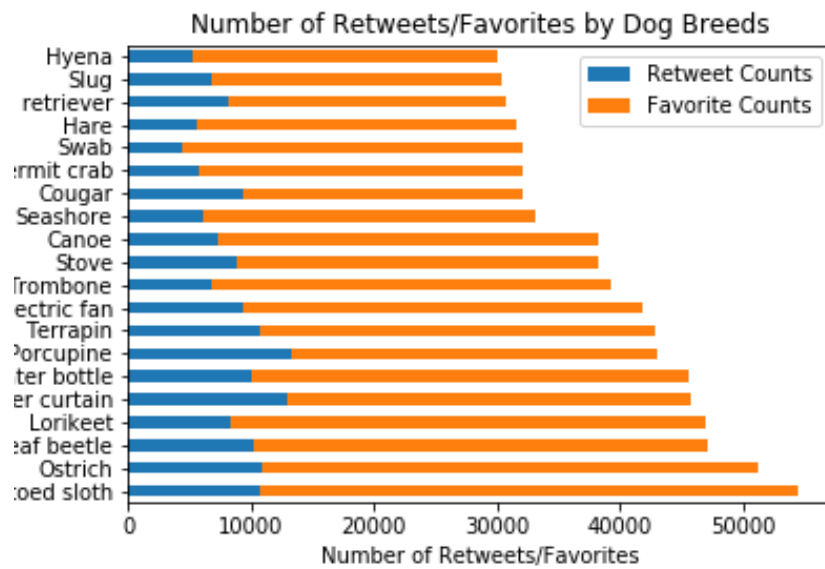
Now, we will see the ratings for different dog stages. In order to enhance the difference, the values of rating will be 10 times larger in the figure. Although the number of retweets/favorites of *puppo* is significantly higher than the other stages, the average ratings are similar to *doggo* and *floofer*. The stage of *pupper* is still the least favorite among the others.



3. Popularity of Different Dog Breeds

There are three dog breed predictions in the dataset, where #1 prediction of breed prediction based on images generally has the higher confidence among all three predictions. Therefore we assume that #1 breed prediction is the actual breed of dog in the tweet.





From the first graph above, we can see that Golden Retriever, Labrador Retriever, Pembroke, Chihuahua, and Pug are the most common dog breed that shows on the tweet of WeRateDogs. However, these dog breeds generally do not have the most retweets and/or favorites among all breeds. None of the top 20 common breeds is on the list of top 20 tweets with most retweets/favorites. It might also be a kind of human behavior that people usually retweet/like a tweet when they see something rare.