WeRateDogs Analysis

Daniel Cummings

7/26/2018

Overview

This analysis is of a tidy dataset consisting of tweet data from [WeRateDogs](#) (Twitter [@dog_rates](#)) and associated dog breed predictions from an image processing deep learning model. The WeRateDogs account gives a rating to people's dogs using a humorous commentary and rating system (e.g. 7/10, 11/10, 13/10, etc.). The focus of this analysis will be on dog ratings and popularity (retweet and favorite counts) and how they relate to dog breed and [DoggoLingo](#).

Analysis

The main features of the WeRateDogs tweets are the rating system and commentary. The distribution of the dog ratings is shown in the following figure where the ratings follow a left-skewed distribution with the median of 11 (over a denominator of 10). In other words, more than half of the ratings are greater than 10/10, which is in the humorous sprit of WeRateDogs.
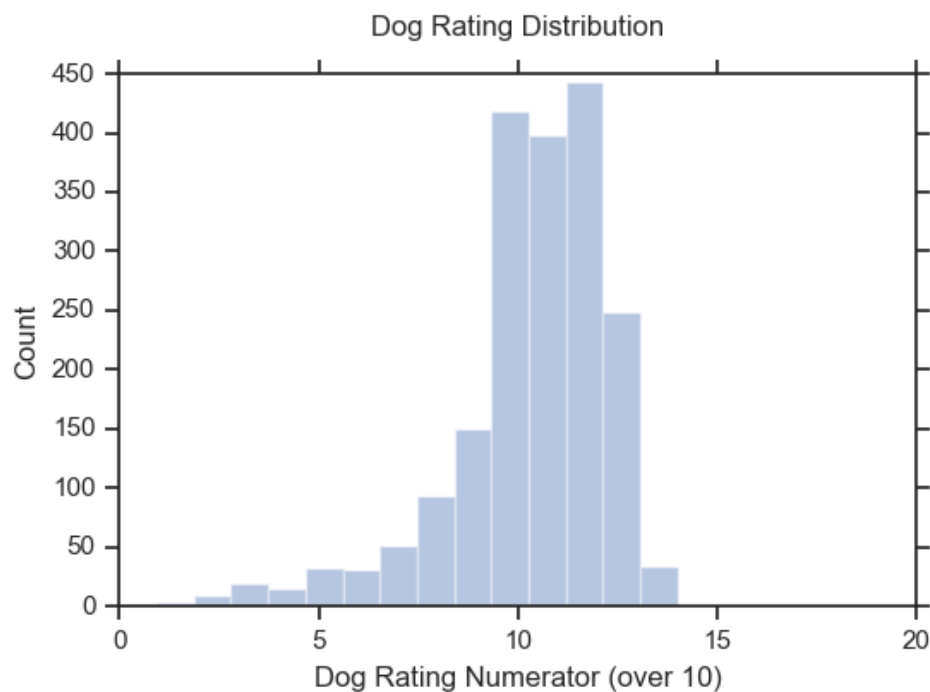
How does the rating relate to the tweet favorite counts? The following figure is a linear regression that shows a positive correlation between rating and favorite count. There are very few cases where a rating of less than 10/10 has more than 20,000 tweet favorites.
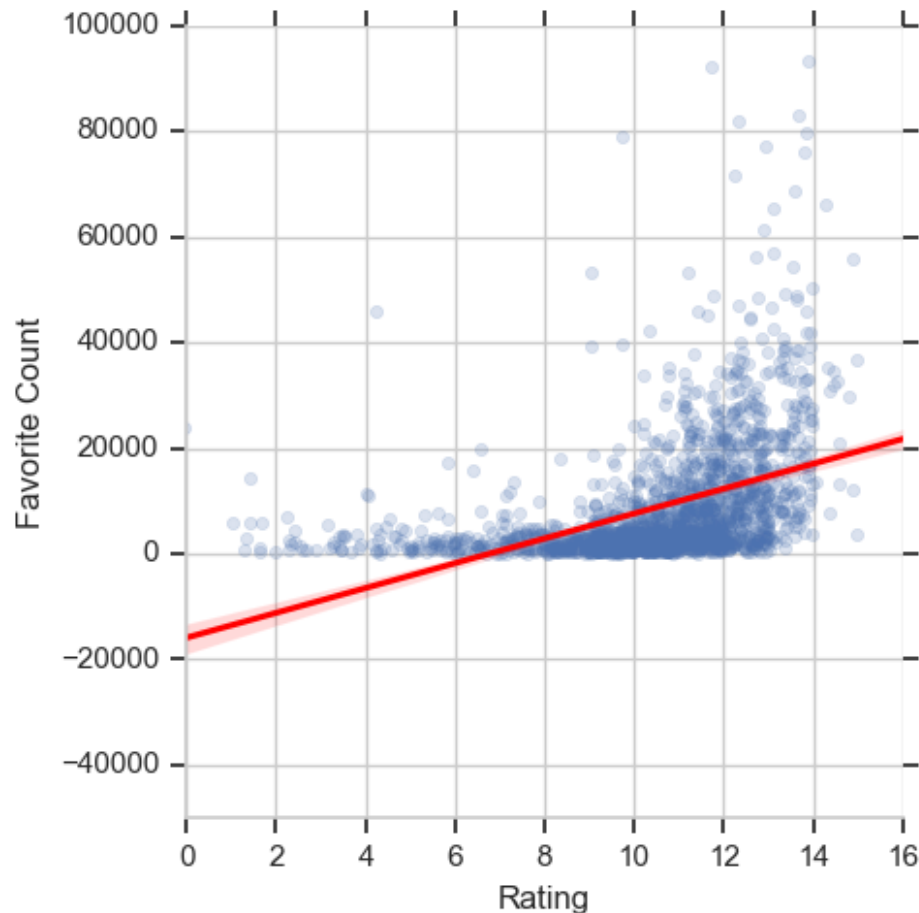

DoggoLingo is pervasive in the WeRateDogs culture. The following list gives a summary of each term that was commonly found in the dataset.

- doggo: a big dog, usually older and more experienced
- pupper: a younger and inexperienced smaller dog
- puppo: dog equivalent of a teenager
- floofer: big happy fluffy dog

From the perspective of ratings and DoggoLingo category, *floofers* and *puppos* receive the highest median and mean ratings.

Rating by Doggo Lingo
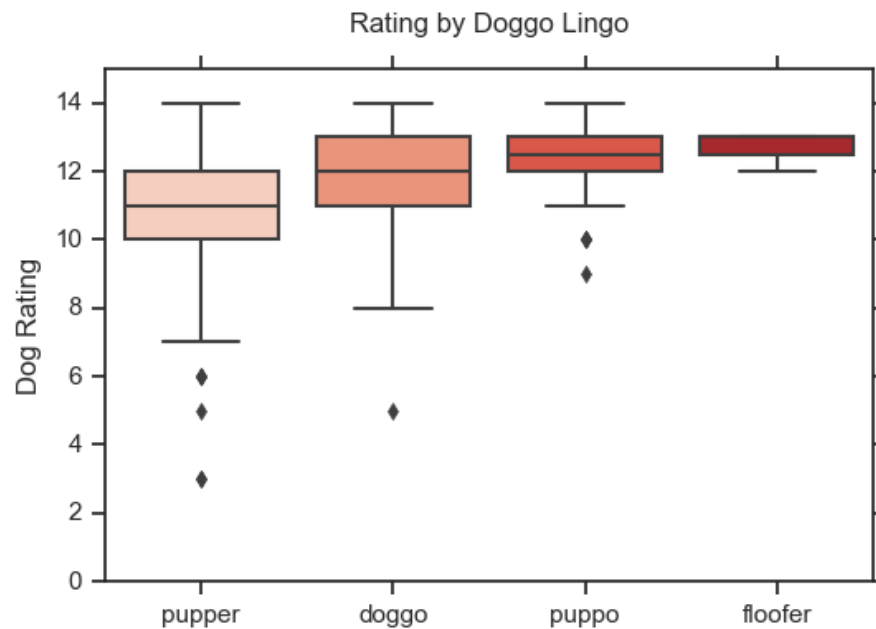
In terms of popularity, *floofer* leads the pack in median and mean retweet/favorites counts. However, the other categories have special cases that exceed the maximum popularity of the *floofer* category.

Popularity by Doggo Lingo
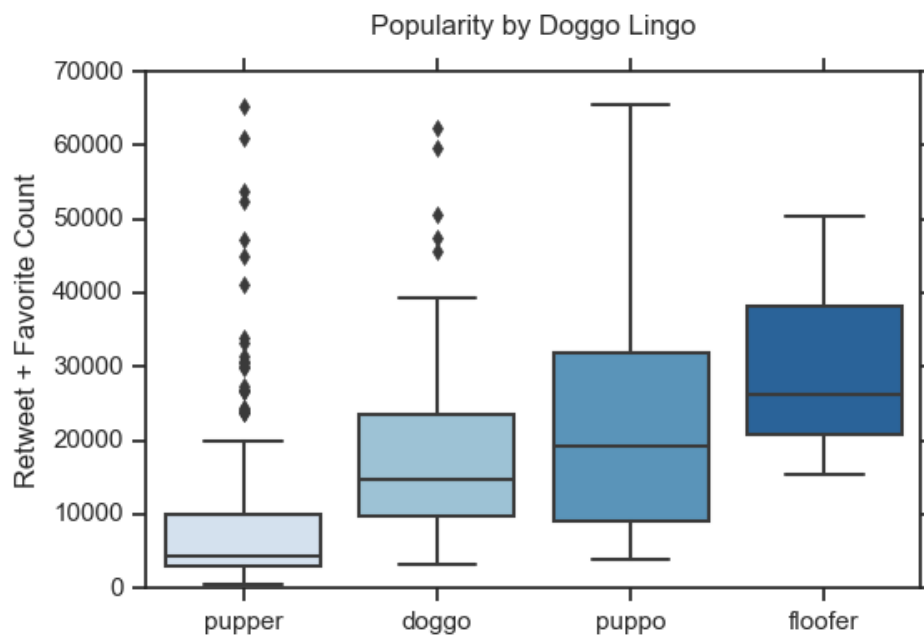
Since DoggoLingo categories are somewhat vague and subjective, it makes sense to look at the top ten most often tweeted breeds. Interestingly, out of the top ten, four of the five larger breeds (Golden Retriever, Labrador, Samoyed, Choy) had a higher median popularity than all of the small breeds (Pug, Chihuahua, Pomeranian, Toy Poodle). Followers of the WeRateDogs Twitter account generally seem more likely to 'favorite' or 'retweet' large *floofers* over small *puppers* and larger breeds over smaller breeds. For future work, it would be interesting to break this down by gender and age group.

Popularity by Breed