

# The Act Report of WeRateDogs Project

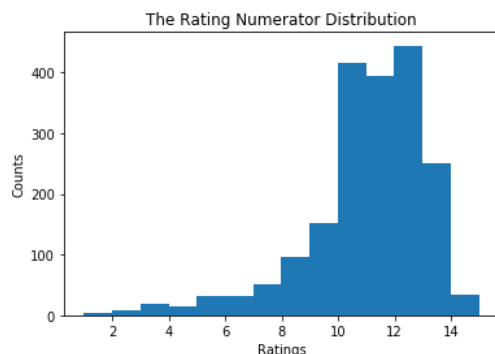
By Z.Jiang  
May 2020

'WeRateDogs' is a Twitter account which rates people's dogs in a humorous way. According to Wikipedia, the WeRateDogs was started in 2015 by Matt Nelson and has received international media coverage for its popularity and some disputes the activity raised. The Twitter account has nearly 4.56 million followers by December 2017.

The WeRateDogs Twitter data provided by the Udacity as part of my Data Analyst Nanodegree program only included tweets up to August 1st, 2017 which does not include the retweets, the neural network identified dog breeds information, or the dog image.

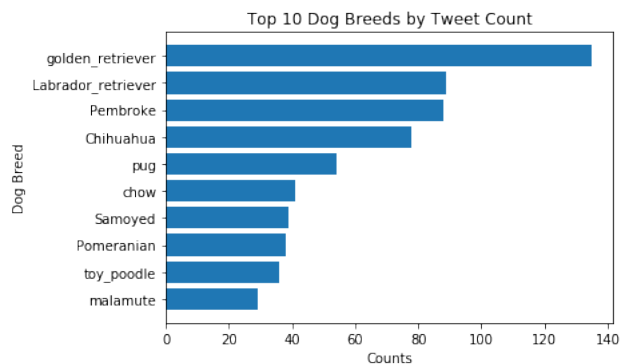
The final dataset contained 2015 observations after the data wrangling procedures.

## Question1: What is the distribution of the ratings numerator?



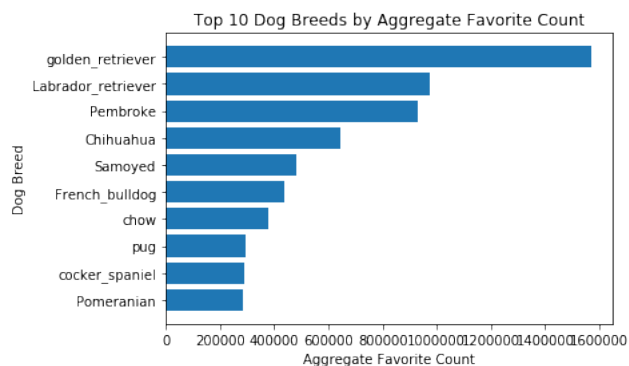
We can see the ratings\_numerator distribution is left skewed and most of the ratings are between 10 and 14. The value of the ratings\_numerator is the number given by the Twitter users showing how much they like the posted dog picture. The counts distribution indicates the Twitter users generally over rate the posted dog pictures.

### Question2: Which dog breeds were most tweeted?



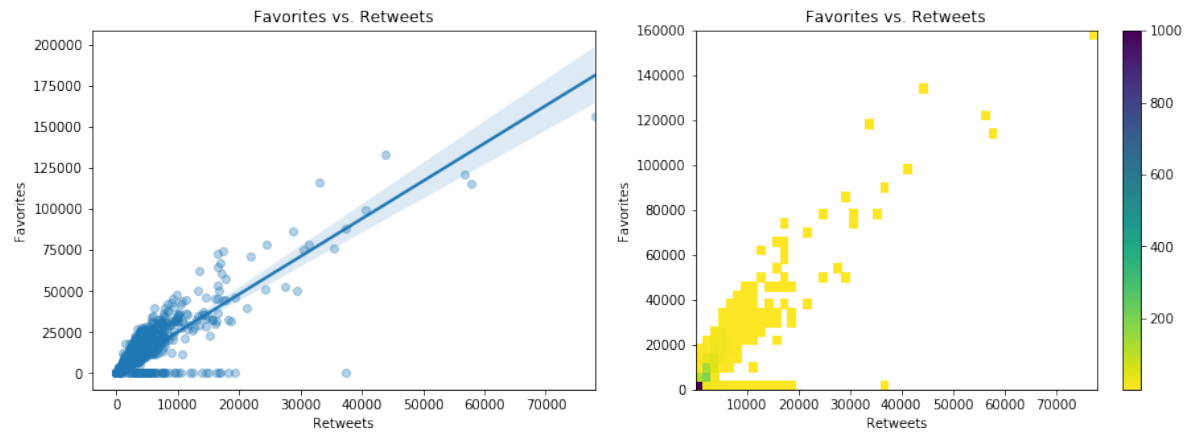
The neural network identified bod breed that receive most tweets is the Golden Retrievers, followed by Labrador Retrievers and Pembrokes.

### Question3: Which dog breeds receive the highest favorite counts?



The pattern of the most favorite dog breeds is very similar to the most tweeted dog breeds pattern.

**Question4: Is there a correlation between the retweet counts and the favorite counts?**



The regression plot, heatmap of 'retweets' and 'favorites' suggest the 'retweets' is positively correlate with 'favorites' and the  $r^2 = 0.93$ , indicating a strong positive relationship between the two variables.