

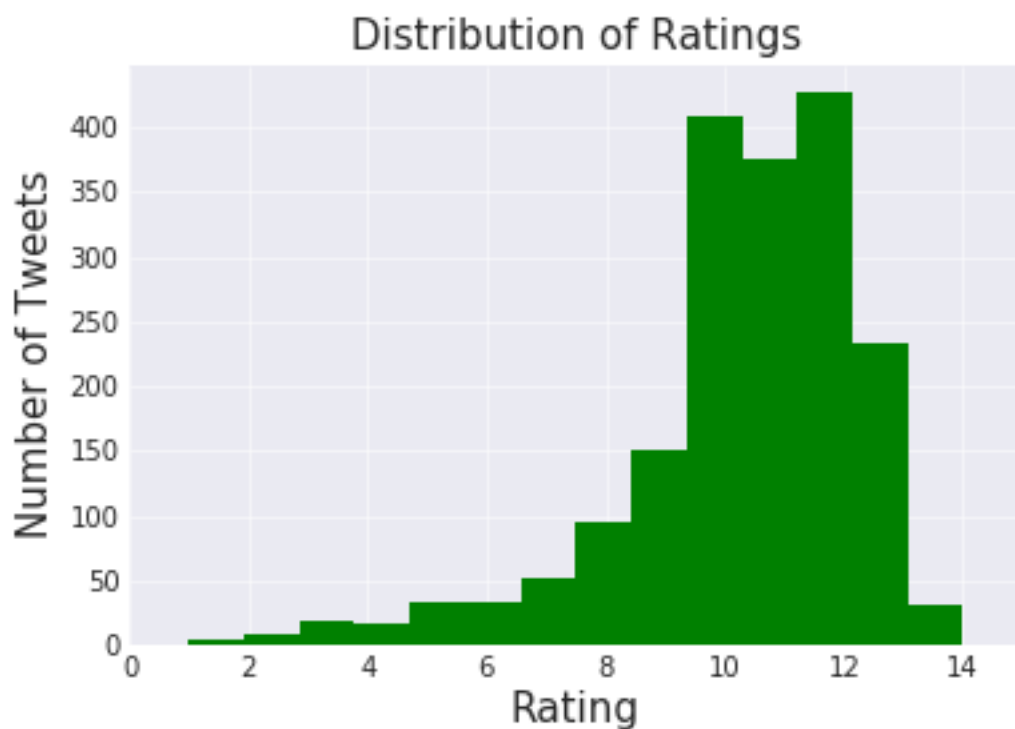
Analysis of WeRateDogs Reports

This report is a quick analysis for different data sets related to WeRateDogs page on twitter,

As of today, WeRateDogs has 8.8M follower on twitter, it gained this publicity by rating peoples dogs with a good sense of humor, with the rating process they sometimes give more than 10/10, so it's normal to see a dog with 14/10 rating (Yes, it's not an error!)

We can start by this point for the first part in our analysis

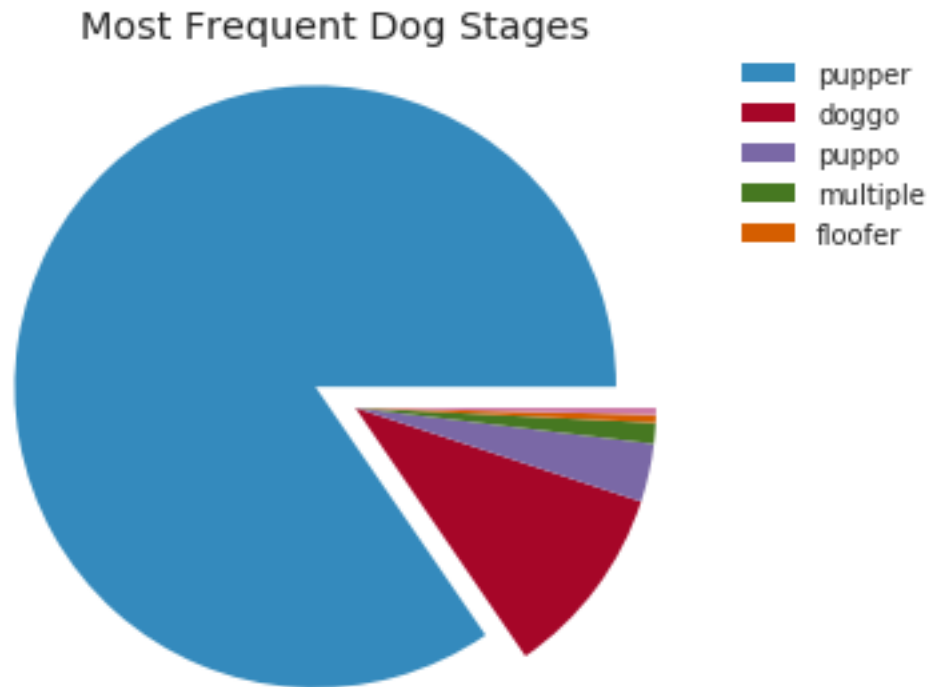
Distribution Rating



As we can see in the distribution rating histogram, data is left skewed, as 75% of rates are above 10/10, which gives the idea that rating is used in funny way more than actual rating from 10

Dogs Stages

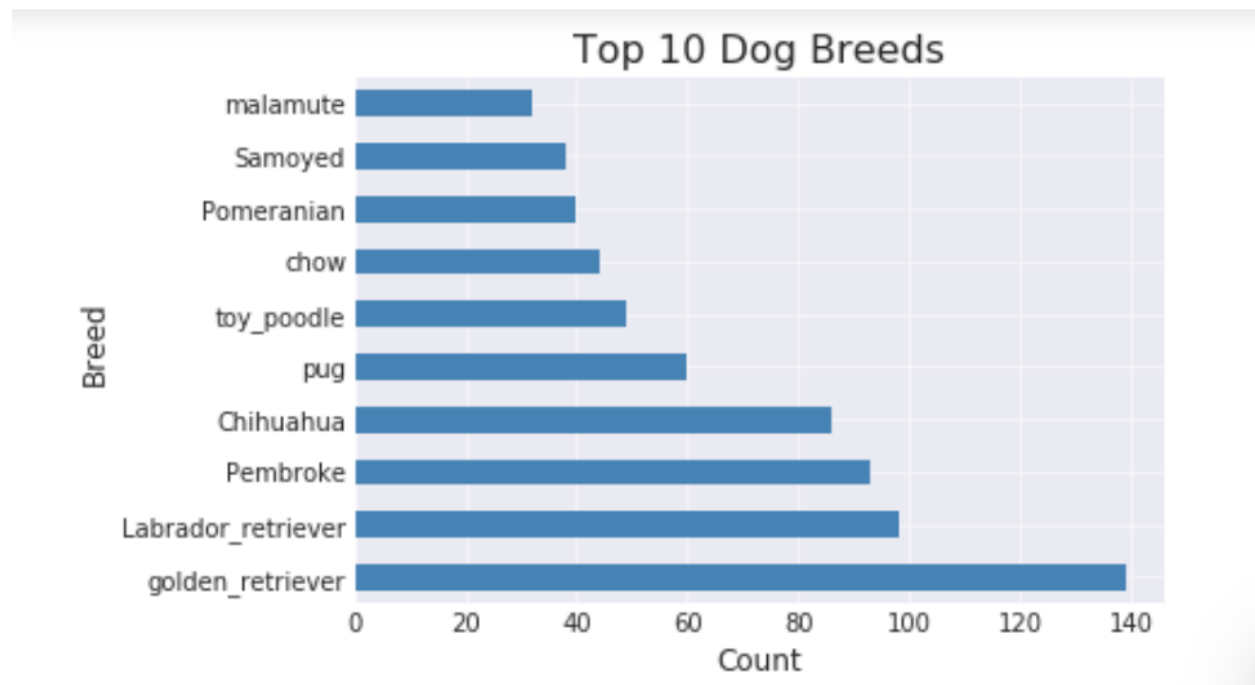
We can have a closer look into the dog stages that has been tweeted in the WeRateDogs in the graph below



It shows that pupper is most common dog stage within the shared dogs, excluding dogs with no stage mentioned of course, and then doggo comes in the second place.

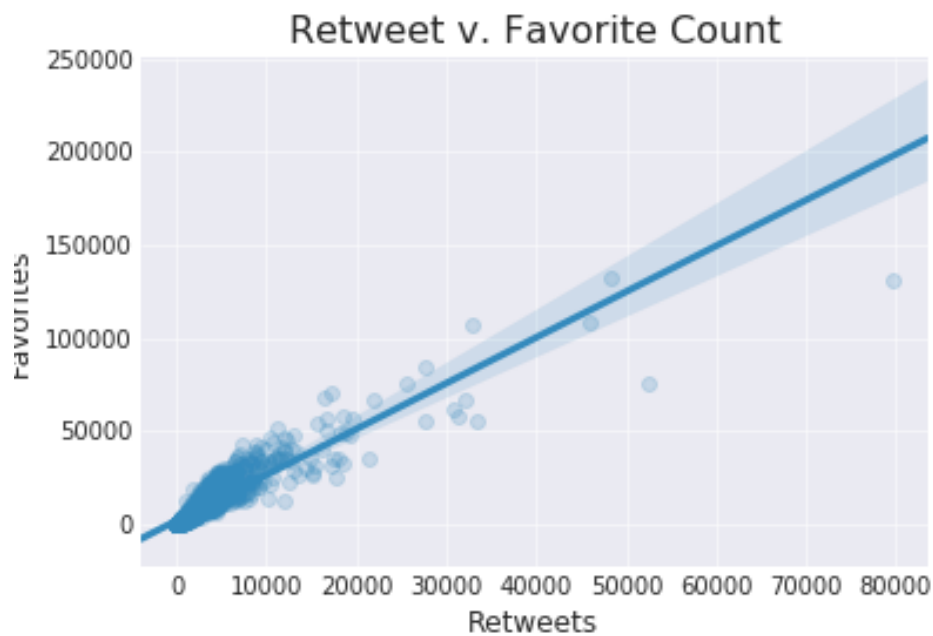
Dog Breeds

Then moving from dog stages to dog breeds, we will find below top 10 dog breeds mentioned in the tweets, worth noticing that golden retriever is much popular than others dog breeds, then comes after it, Labrador retriever, Pembroke, Chihuahua



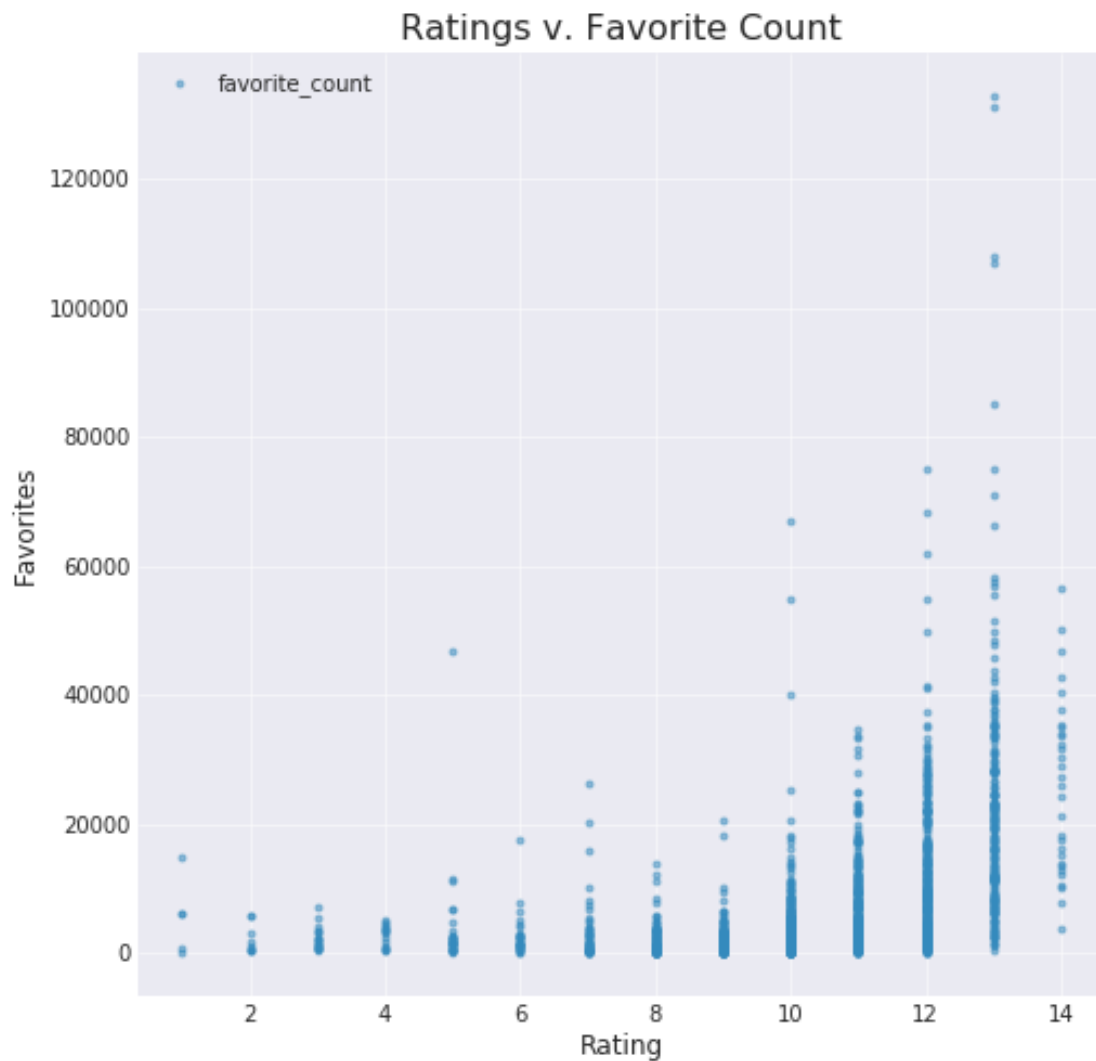
Favorites, Retweets relationship

The graph below presents very interesting points, which is the relation between favorites and retweets, we can notice that there is a strong positive relationship between both retweets and favorites, the more tweet has favorite the more it's retweeted, and this is logical, because it means the tweet is loved by many people, and this is presented in the way of reaction to the post



Favorites, rating relationship

Same concept as before, the positive relationship between rating and favorite, but this time it's a little bit different, because in the previous graph, both actions were done by the viewer, this time one action is from the page side, which indicates, that the rating given by the page, leads the favorite count in a positive way.

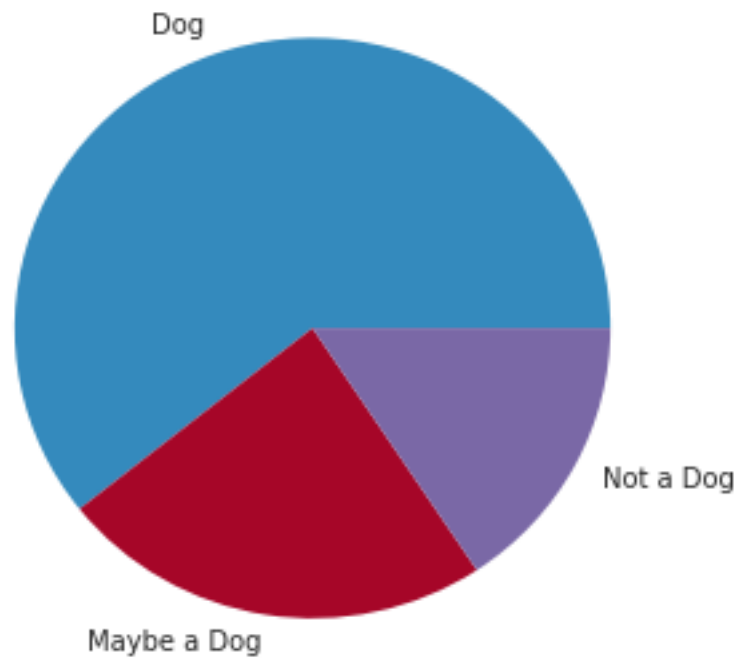


Dog Image Prediction

This graph represents the percentage of dogs which were identified by machine learning software, to determine the type of the dog.

Around 60% of the dogs were identified as dogs by the software, while another 23% might be dogs, the rest could not be identified as dogs at all, which gives us insight about the development of machine learning in this field, 60% is a good percentage, but still not good enough to be afraid from machine revolution against humanity.

Dog Image Predictions



Finally

Difference between data accuracy before and after data wrangling is really amazing, which explain the idea of the importance of data wrangling.