

Insight 1

As the value_counts method suggested, golden retriever is identified as the most popular species among dog. In addition, Labrador retriever and Pembroke also take up a significant proportion of the image predictions.



Personally, there are two main reasons to potentially explain why golden retriever as well as the other two species mentioned above would be top of the list: first reason is perhaps these dogs are indeed popular in the community, and people just love sharing their pictures; yet the second reason would be, these dogs have a higher predicted rate in the original image prediction, which

means they were either easier to predict or predicted with a higher confidence level or both. I would argue that both reasons are prominent in the outcome, because the sample size would be large enough to suggest that these dogs are popular, and admittedly, the bar I set for confidence level was not high.

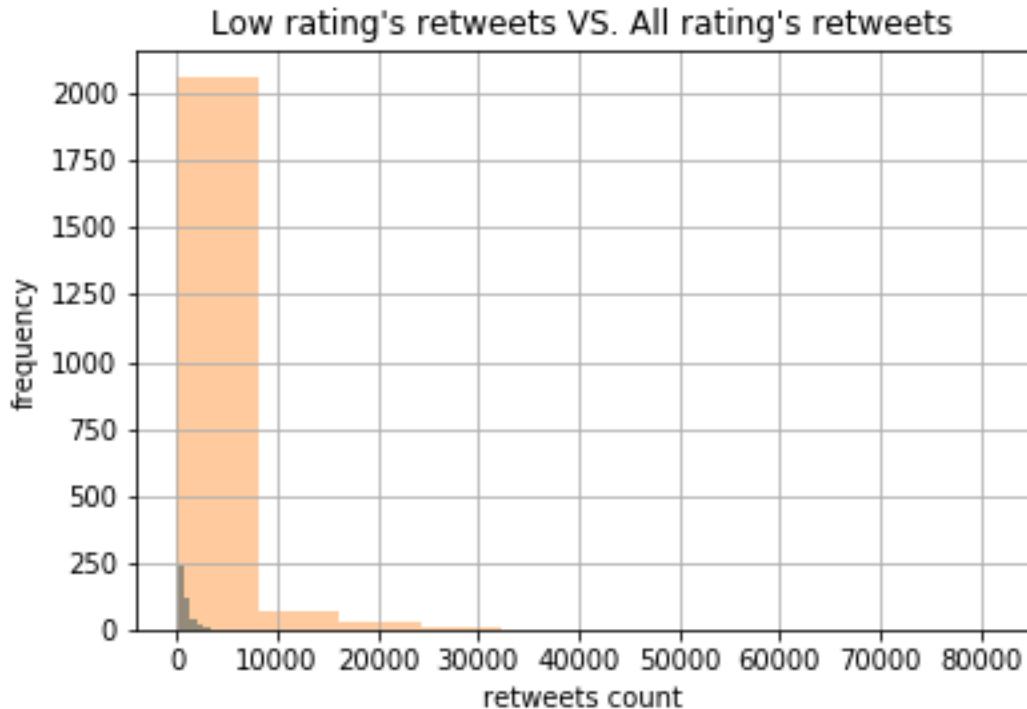
Insight 2

Pupper is the most popular dog stage among four, and it is way much more than the others. Why is that? Well small dogs are arguably cuter and more touchable to humans, and so people love to send more pupper to the community. Noting that values of the dog stage' were extracted from the tweets' text, and so it was the Twitter people themselves who labeled their dogs as pupper, doggo and etc. Be mindful that there is no rigorous way of categorizing the 4 'dog stages' now, but will there be a more standard yet innovative and cute way to identify a dog's stage in the future? We will see.

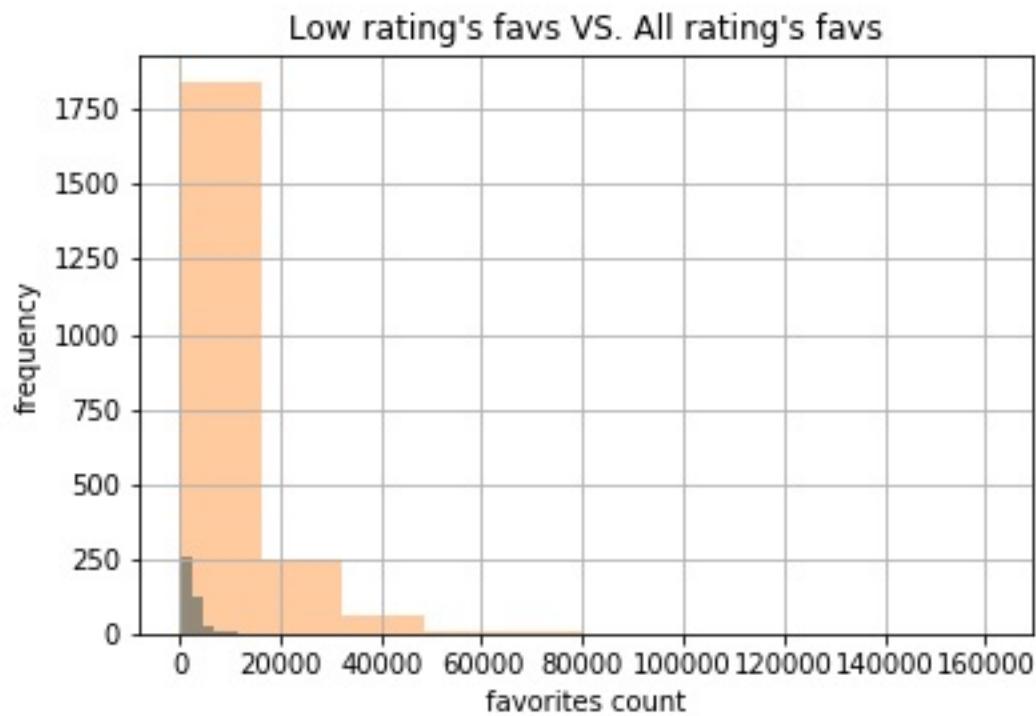


Insight 3 and Visualization 1

What if people rate their dogs low? We all know this is uncommon in the WeRateDogs community, as the ratings seem always above 10. However, there are still some tweets with a lower than 10 rating in this dataset. Take a look at the two charts:



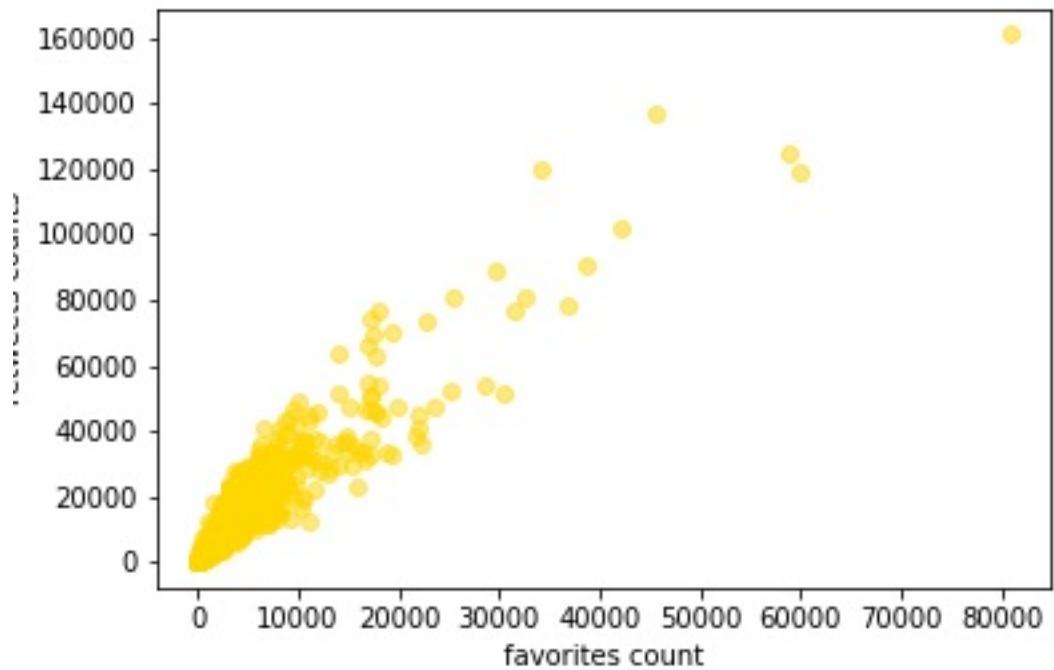
Low rating's retweets VS. All rating's retweets, the chart compares the retweets count of the low rating group: who rated their dogs lower than 10; and the total group. Seeing the little dark area on the left bottom? That's how many retweets people got when they rated their dogs low.



Low rating's favs VS. All rating's favs, the same goes to favorites as well, few people chose to favorite a low rating tweet.

Insight 4 and Visualization 2

Generally, if a tweet had many retweets, then it should gain many favorites as well. This is illustrated in the scatter plot below:



Favorite count VS. Retweets Count, there is an evident positive correlation between retweet and favorite in the scatter plot above. If people retweeted a tweet, they probably wouldn't be stingy to also give it a favorite.

