A Dog is a Dog - is a Dog?

by Petra Schneckenburger

#WeRateDogs is a well-known site on twitter. Users can send in pictures or videos of their dogs, they are rated by the team and posted on the site where everyone can comment, like or retweet it.

This site was object of a project of Udacity's Data Analyst Nanodegree program. We students have been given part of the archive of WeRateDogs, together with prediction values of a neural network model trained to recognize dog breeds in images. The task was, to bring all given data in a clean and tidy state, so that the data easily can be analyzed (so called Wrangling) and to perform some analyzes on them.

All the wrangling efforts resulted in a dataset based on 1968 tweets, and here are some of my findings:

The Ratings

Ratings were usually with a denominator of 10, but a numerator of more than 10, so 11/10, 12/10 and so on. There are numerators less than 10, especially when pics handed in contained other animals than dogs. Sometimes they deviate from denominator 10, mostly when there is a bunch of dogs seen, calculating the ratio makes these fit in again. The rating ratio in the examined data was from 0.1 to 1.4

Who is the best dog? Meet Atticus!

There is one dog with an outstanding rating of 1776/10.
Just besides, 1776 was the year of the United States Declaration of Independence

https://t.co/GRXwMxLBkh

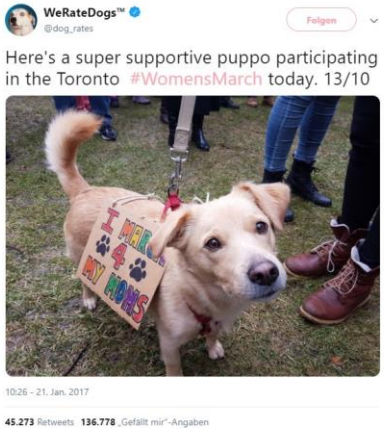
One of the least rated 'dogs'…

… obviously isn't a dog.

There were several other animals in the tweets nicely commented by the team.
And I really don't understand the low rating.

Favorite and Retweet Counts

Since the rating is done by a small team and thus can be very subjective, the amount of likes (here: favorites) and retweets might also be good criteria for evaluation.

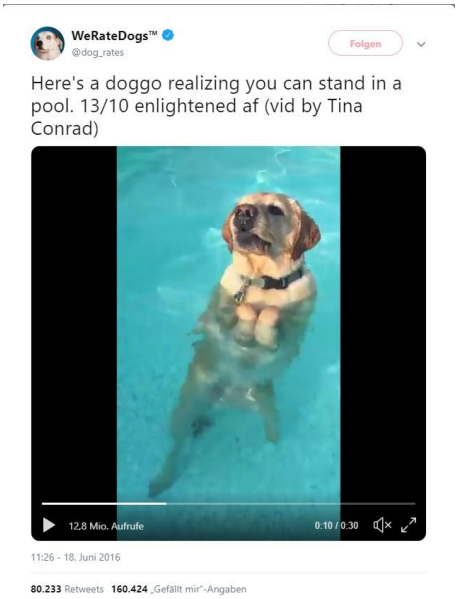


Left the dog with most favorite counts

and on the right the one with the most retweet counts

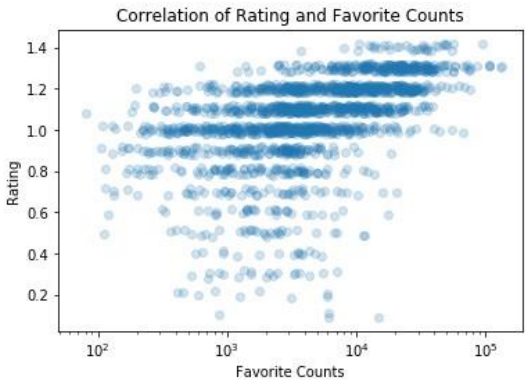
https://t.co/nTz3FtorBc

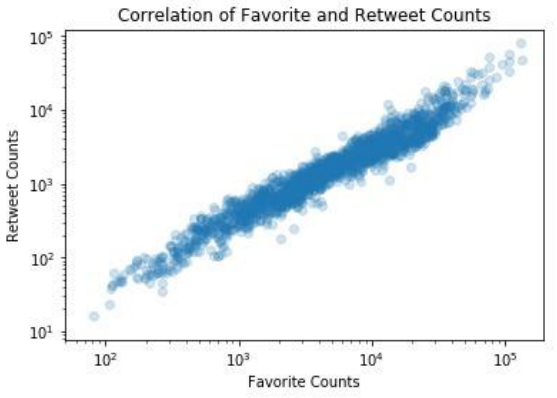
https://t.co/7wE9LTEXC4



Interestingly the least favored and least retweeted picture was the same: https://t.co/11LvqN4WLq

How are ratings, favorite counts and retweet counts related?

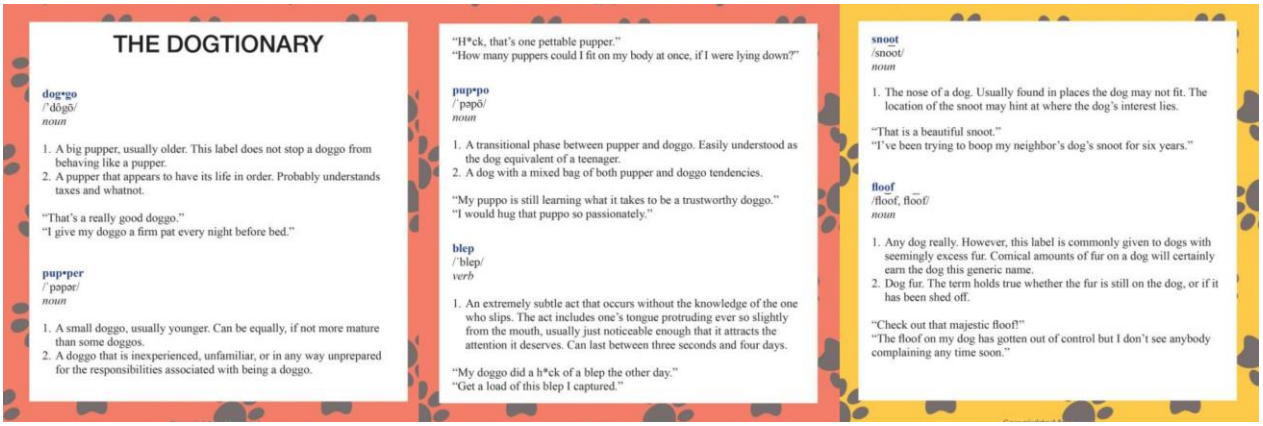




Favorite and retweet counts seem highly positive correlated, this makes sense, the more people see it the more can like it. Interestingly the correlation between the given rating and favorite count is not so clear, only seen on higher ratings. This may be due to the fact that people also tend to like things that are uncommon or weird, as non-dog pictures on a dog site. In case of dog pictures nevertheless, favorite counts can be used as a measure of maybe less subjective user ratings.

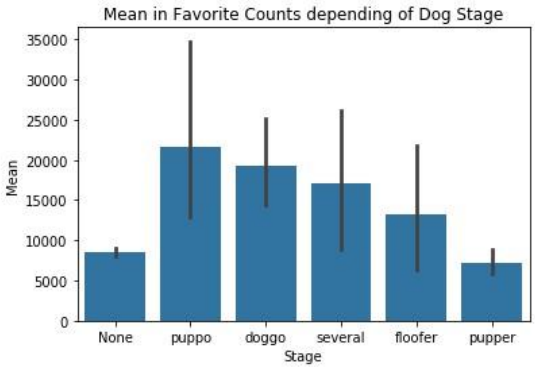
Another speciality: Dog stages

On WeRateDogs there are some very special terms for dogs, relating to the age or the furriness. Details are shown below.



The Dogtionary explains the various stages of dog: doggo, pupper, puppo, and floof(er) (via the #WeRateDogs book on Amazon)

On part of the twitter data were these stages given. I was curious, if some stages were preferred to others.

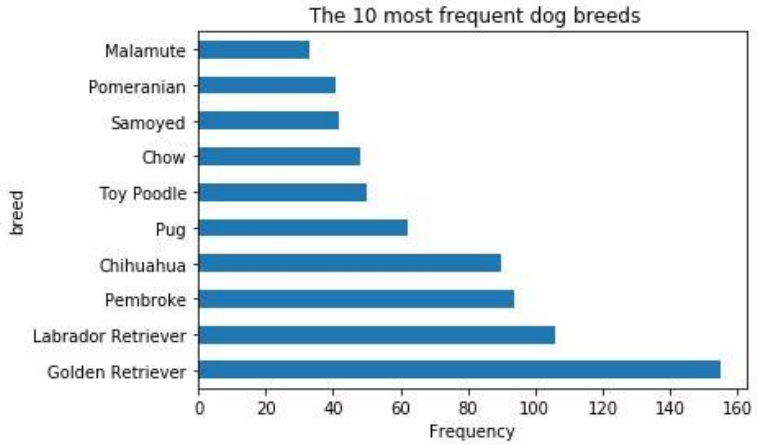


Interestingly, having a dog stage indicated in the text, seems to lead to much more favorite counts – with the exception of pupper, that has similar favorite count mean than the not staged dogs. This is surprising, since pupper is the youngest stage, I would expect more likes due to natural sweetness of young dogs. The limitation of these data however is, that the group sizes are small and therefor the variance is large, as can be seen by the ugly error bars indicating standard deviation. On higher amount of stated dogs, the seen effect may disappear.

The dog breed predictions

In the data given, the dog prediction algorithm found 113 different dog breeds.

Golden Retriever is the most common dog breed found in the dataset, followed by Labrador Retriever. Interestingly the next dog breeds are much smaller in size. Seems there is no strong preference



to large sized dogs by WeRateDogs users, but it would be interesting to examine this further.

Limitations of the breed prediction

Out of the analyzed dataset, 303 pictures were not identified as dogs. Some of them in fact were other animals, or even things like a fan, but there were many dogs missed, so the true ranking in frequency may differ from the shown above.

Together with the dog breeds, the prediction algorithm gave a certain confidence, that the prediction is true. In the raw dataset, there was one picture with a confidence of 1, but not the dog was recognized, it was (correctly) addressed to the jigsaw puzzle in the foreground.

The model also had problems to recognize dogs in shopping carts, in front of TV and similar situations.

However, this sweet dog on the right was recognized correctly, with a confidence of 0.00001!

I am not experienced with image prediction models, but I think the model used here did a really god job.