An insight glimpse on #WeRateDogs

# Introduction

Through Data wrangling I managed to extract a clean dataset that shows the useful information of the users’ interaction to the popular hashtag.

I found 8 quality issues and 2 tidiness ones that I cleaned and republished my work. This document is to show the useful insights I managed to get.

# Analysis

## How well did the image predictor do?

So #WeRateDogs was passed by an image predictor machine to predict the dog breed per each picture.

### Dog image prediction’s average confidence percentage (total)

1. The average confidence percentage with the failed predictions is 42.60%.
2. The average confidence percentage without the failed predictions is 54.95%.
3. These failed predictions represent 22.48% of the total tweets in this dataset.

Showing the top 10 average confidence percentage the image predictor got, here we can see that the top breed that has been detected with a higher accuracy is the komondor (97.25%).

|  |  |
| --- | --- |
| prediction | confidence level (in decimals) |
| komondor | 0.972531 |
| clumber | 0.946718 |
| keeshond | 0.844431 |
| bull\_mastiff | 0.815618 |
| brittany\_spaniel | 0.775591 |
| bernese\_mountain\_dog | 0.766685 |
| samoyed | 0.71744 |

On the other hand, the Irish wolfhound breed (6.31%) was the least image predicted in terms of average confidence percentage as shown in the 10 least breeds in terms of confidence percentage.

## Who is twitter’s popular dog?

The popularity contest in Twitter can be seen by either the number of retweets for a post or the most favoured tweet of all.

### The most favored dog to twitter's heart is...



Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10 https://t.co/nTz3FtorBc, unfortunately we could not find a name to our beloved dog. With 132,810 favorite tweets, this unnamed dog captured the hearts of many users.

Ironically that beloved to twitter society dog was breeded wrong as Lakeland Terrier, with confidence percentage as ~20%. Also, this tweet post did not include that puppo’s name.

### The most dog retweeted

Now this popular doggo's image link is broken, so cannot be displayed ☹. This unnamed dog's image was predicted as a Labrador retriever breed, with ~82.53% confidence percentage, a better return than the most favored puppo up there.

## Which dog should you buy?

If you are looking to have dog and go popular on your Twitter, I gathered to you these common breeds, now since each person would have their own taste in choosing a dog, this table right here combines both the top 10 most retweeted and favorited breeds.

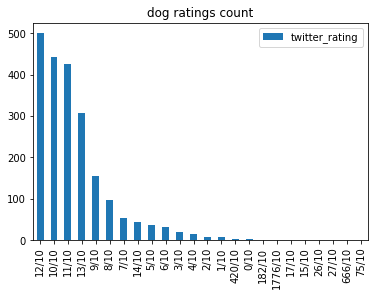
|  |  |  |  |
| --- | --- | --- | --- |
| ***prediction*** | ***Retweet\_count*** | ***prediction*** | ***Favorite\_count*** |
| bedlington\_terrier | 7510 | saluki | 24060 |
| afghan\_hound | 5976 | bedlington\_terrier | 21153 |
| standard\_poodle | 5508 | french\_bulldog | 18416 |
| english\_springer | 5401 | bouvier\_des\_flandres | 18032 |
| saluki | 5133 | afghan\_hound | 17326 |
| french\_bulldog | 5039 | black-and-tan\_coonhound | 17012 |
| cocker\_spaniel | 4858 | flat-coated\_retriever | 16791 |
| samoyed | 4843 | irish\_water\_spaniel | 16400 |
| whippet | 4840 | leonberg | 14934 |
| great\_pyrenees | 4784 | english\_springer | 14138 |

I marked to you the breeds that will mostly likely get you the most retweets and favorites if you plan on posting 😉.

* afghan hound
* bedlington terrier
* french bulldog
* saluki

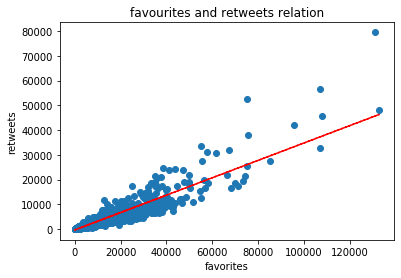
Other breeds will do fine, they will guarantee you social media popularity. Of these, I personally vote for the [saluki](https://en.wikipedia.org/wiki/Saluki).

## How generous is @dog\_rates in their ratings



Although we have some overhyped ratings, they do not represent anything but one record each (and they are around 4 records only). But in general, twitter ratings show a generous trend towards the 10/10, 11/10, 12/10, 13/10.

## Is there a corelation between the retweets and favorite counts?



There is a positive correlation between the favorited tweets and retweets. It seems that the @dog\_rates are achieving a high user traffic.

# Conclusions

We can summarize the conclusions to these points

1. The image prediction neural network is not accurate enough, as nearly 484 records were not identified/identified properly as it can be seen from the averages.
2. A doggo and a puppo own Twitters #WeRateDogs scene.
3. Of the average breeds’ favorite and retweets, commonly you can have a Saluki or any other of these dog type and they will make you famous on Twitter.
4. @dog\_rates showed their generosity in their ratings as they almost always dive a higher rating. And yes, it seems that ['they're good dogs Brent'](http://knowyourmeme.com/memes/theyre-good-dogs-brent).
5. In terms of trafficking, the #WeRateDogs got a good recognition on Twitter’s platform for tweets reaching the 20,000 favorites and retweets marks on average.

Other mentions that need to be addressed is that a heavy number of the names of dogs was not included (or cleaned as ‘None’ as they got invalid names), it is a shame that we could not know the name of these most popular dogs (would have helped you after choosing your dog breed).