

WeRateDogs analysis and insights report

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. WeRateDogs has over 4 million followers and has received international media coverage. After cleaning the dataset almost 1700 tweets were ready to be analyzed.

The entire data set discussed 4 dog types : doggo, pupper, floffer and puppo but which was mentioned the most ? The figure below shows dog class pupper with the highest count of 164 and floffer with the lowest count of 1.

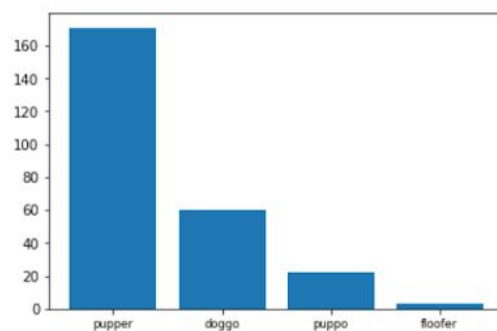


Figure shows count of each dog class in the dataset.

There are several dog classes also known as breeds mentioned in the tweets. However, which breed/breeds had the highest ratings among all tweets collected? Statistics show that the Golden retriever had the highest score of 150 tweets followed by the Labrador retriever with 100 tweets and Pembroke with 88 tweets.

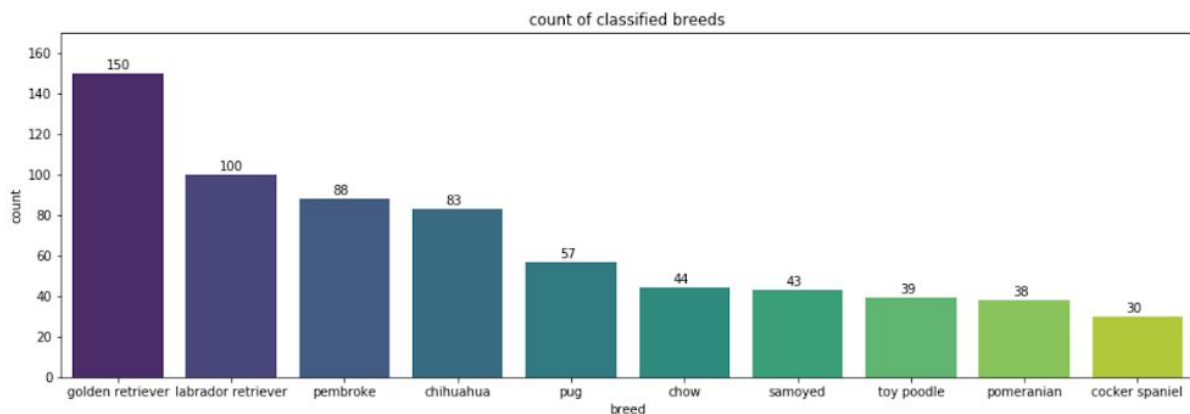
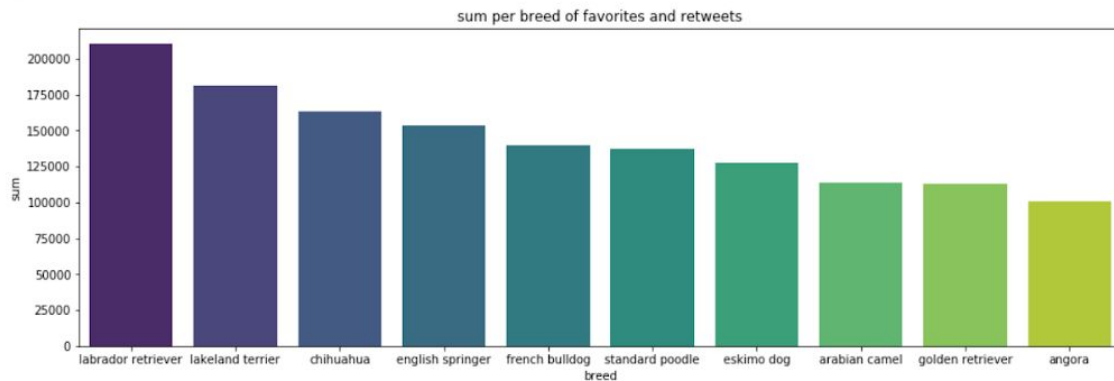
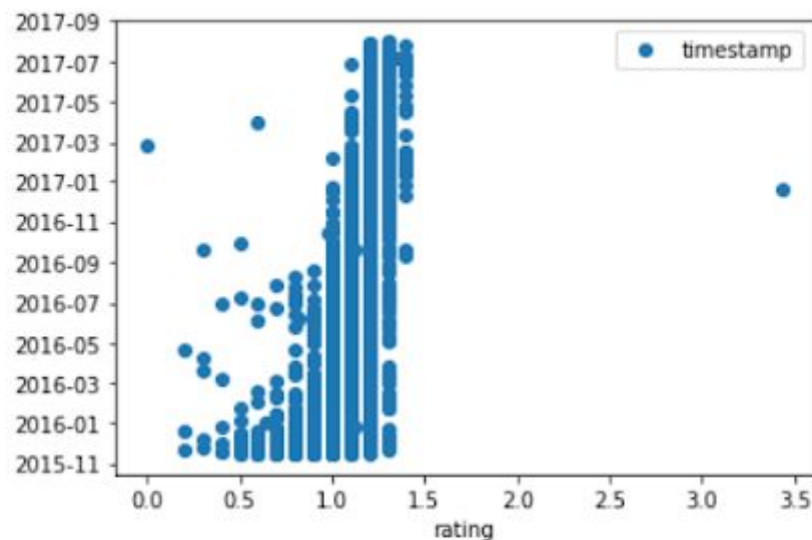


Figure shows top 10 classified breeds with their score count.

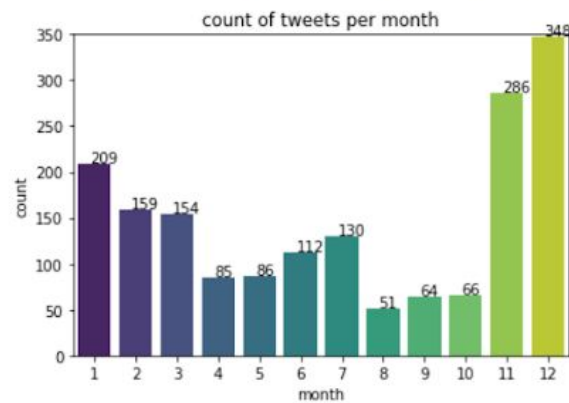
The dataset also held the count of retweets and favorites for each tweet. So here are the top 10 dog breeds with the sum of their retweets and favorites combined. It is clearly stated that the Labrador retriever has the most retweet and favorite count.



To test if throughout the years the ratings differed or not. A scatter plot was made to investigate this point. By the end of 2015 the ratings were between 0.2 and 1.4. As time went by reaching 2017 the ratings were between 1 and 1.5 indicating a higher average rating.



Using the date information in the dataset, we were able to count each tweet per month. December was the best month with tweets with a score of 348 tweets.



Descriptive analysis of the ratings column which was extracted from the tweet's text.

```
In [82]: df_merged_clean['rating'].describe()
```

```
Out[82]: count    1750.000000
         mean      1.088035
         std       0.186992
         min       0.000000
         25%       1.000000
         50%       1.100000
         75%       1.200000
         max       3.428571
         Name: rating, dtype: float64
```