

Act Report

Cyril Ocloo

July 7, 2022

1 We Rate Dogs Act Report

1.1 Introduction

The goal of this report is to analyzing, visualize and give insight into the gathered, assessed and cleaned data.

The dataset is the tweet archive of Twitter user @dog_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

The cleaned dataset contains 1299 observations. Pertaining to my insight, the main attributes considered for this report include dog rating, dog stage/cycle, favorite counts and retweet counts.

1.2 Visualizations and Insights

- Here are some general basic statistics pertaining to some of the attributes mentioned in the introduction.

	dog rating	favorite count	retweet count
count	1299	1299	1299
mean	12.172841	8358.274827	2581.046189
std	50.322282	11176.670627	3999.477910
min	1.000000	81.000000	16.000000
25%	10.000000	1762.000000	603.500000
50%	11.000000	3904.000000	1309.000000
75%	12.000000	10469.000000	3107.500000
max	1776.000000	107015.000000	56625.000000

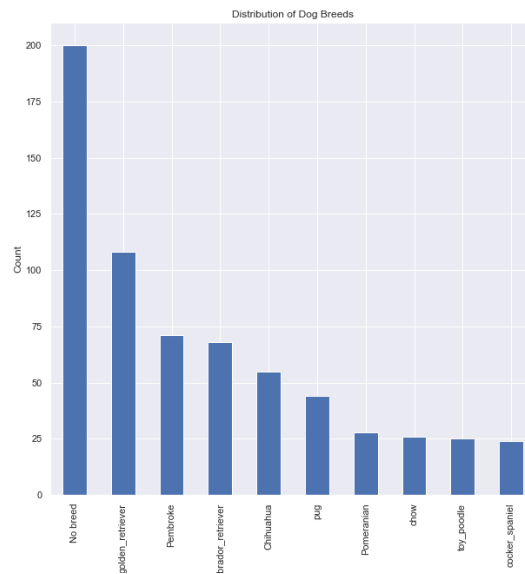
It appears the approximate average dog rating is 12. This includes some outliers that need to be removed or dropped. The average favorite count is approximately 8358 whereas the average retweet count is approximately 2581.

1.2.1 Chart 1: Dog name and their popularity



The dog name Oliver seem to be the most popular dog name. Above is the image of a dog named Oliver and this explains why it is the most popular.

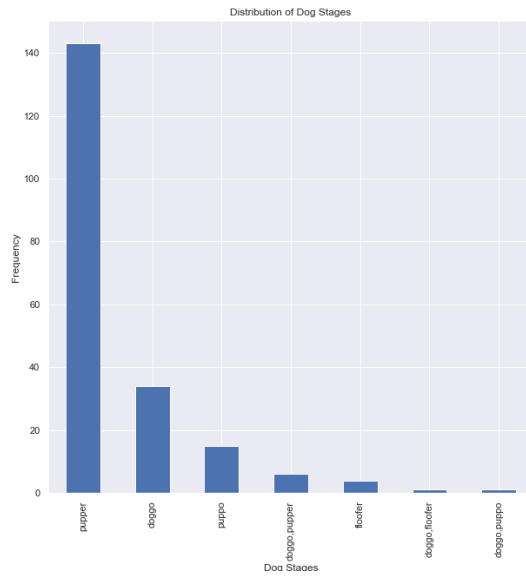
1.2.2 Chart 2: Dog breed and their popularity



It appears about 200 dogs have no breed which is unclear and need to be investigated. However, golden retriever breed happens to be the most common known breed of dogs in our dataset. Image of a golden retriever is displayed above.

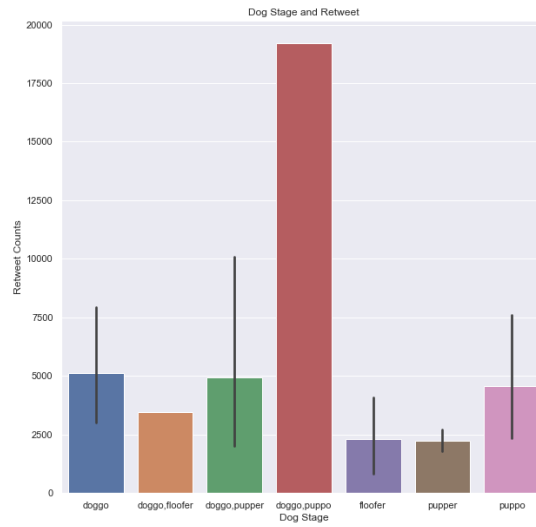
1.2.3 Chart 3: Dog stage and their popularity

Dog stage pupper whose image is displayed below happens to be the most common dog stage whereas doggo, puppo is the the least common dog stage.

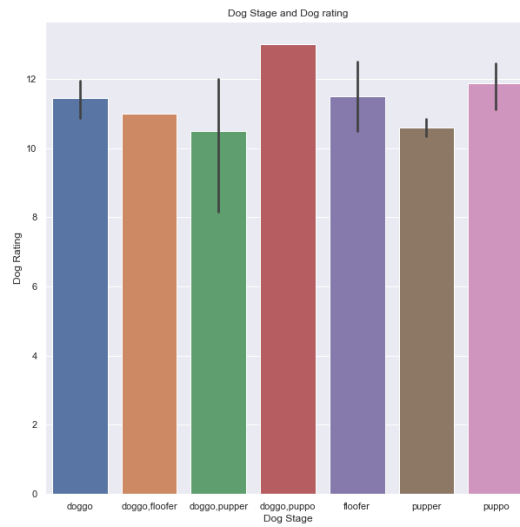


1.2.4 Chart 4: Dog stage and retweet count

From the visualization the least common dog stage/cycle doggo,puppo happen to have the most retweet count and pupper records the lowest retweet count even though pupper is the most common dog stage.

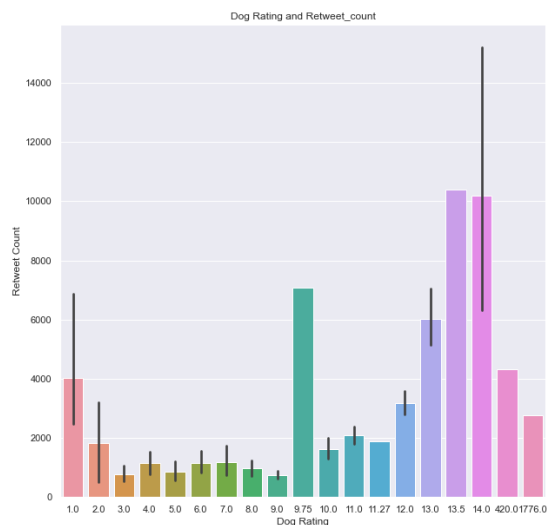


1.2.5 Chart 5: Dog stage and rating



Doggo,puppo records the highest dog rating among the other dog stages and pupper records the lowest rating even though it is the most common dog stage.

1.2.6 Chart 6: Dog rating and retweet count



Obviously, dog ratings of 13.5 out of 10 records the highest retweet count and even the highest favorite counts. This implies that twitter users are more likely to retweet or like a dog with a higher ratings.

1.3 Limitations

- The dataset has some outlier entries in the dog rating which needs to be investigated before removed and will require more time.
- Not all images are dogs even with the predicted confidence, there are still some images misclassified. This can affect the analysis.