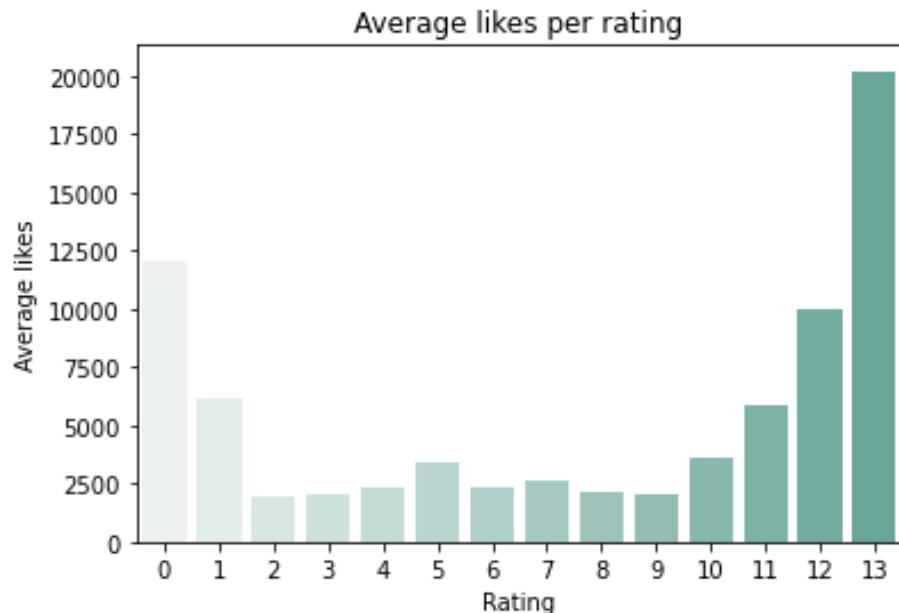


# Act Report

The main object of study in this report is the Twitter account @WeRateDogs, which is an account focused on rating (most of the time) dog pictures, using an unique rating system that usually is based on a numerator greater than 10 and a denominator equal 10. Although it seems to be a simple task, to properly analyse the tweets and ratings from this account it needed some data gathering, assessments and cleaning, which the final result was a dataset with 2057 tweets, from 15/11/2015 to 01/08/2017. Some insights and visualizations are presented below.

## 1. Likes Tendency

We can see a clear tendency of highly rated dogs having also a higher average of likes. This is an expected tendency, as the cuter the dog is, the higher is his rating, and also the number of likes gotten.



The exceptions are the posts with rating 0 and 1, as they are usually jokes or other kinds of posts that are not about rating dogs. Although they receive a great quantity of likes (the rating 0 is the second highest one), they should not be considered

here as we are only interested in dog ratings (the high average is also explained by the low amount of posts in low rating categories).

Not familiar with this breed. No tail (weird). Only 2 legs.  
Doesn't bark. Surprisingly quick. Shits eggs. 1/10

Traduzir Tweet



4:02 AM · 16 de nov de 2015 · Twitter for iPhone

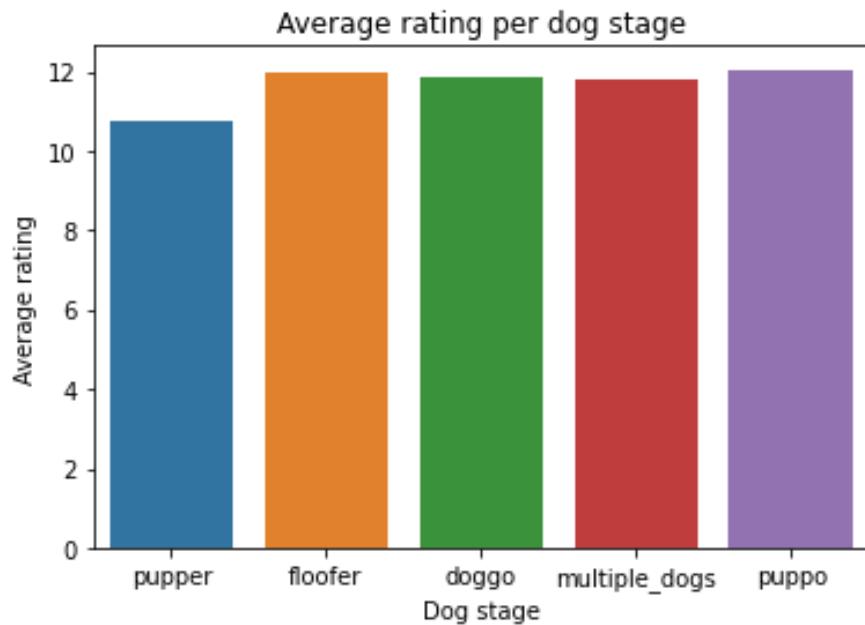
---

5.580 Retweets 118 Tweets com comentário 13 mil Curtidas

*Example of joke (not dog) with rating 1 to be disregarded.*

## 2. Average Rating per Dog Stage

Also regarding the number of likes, we can see how the behavior is when it comes to the dog stage. Unfortunately, only 303 dogs were classified by its dog stage: doggo, floofer, pupper, puppo and multiple dogs (when there is more than one stage listed), making the insights not available for all the dataset. On average all the dog stages are very close, varying between 10.7 and 12 of rating.



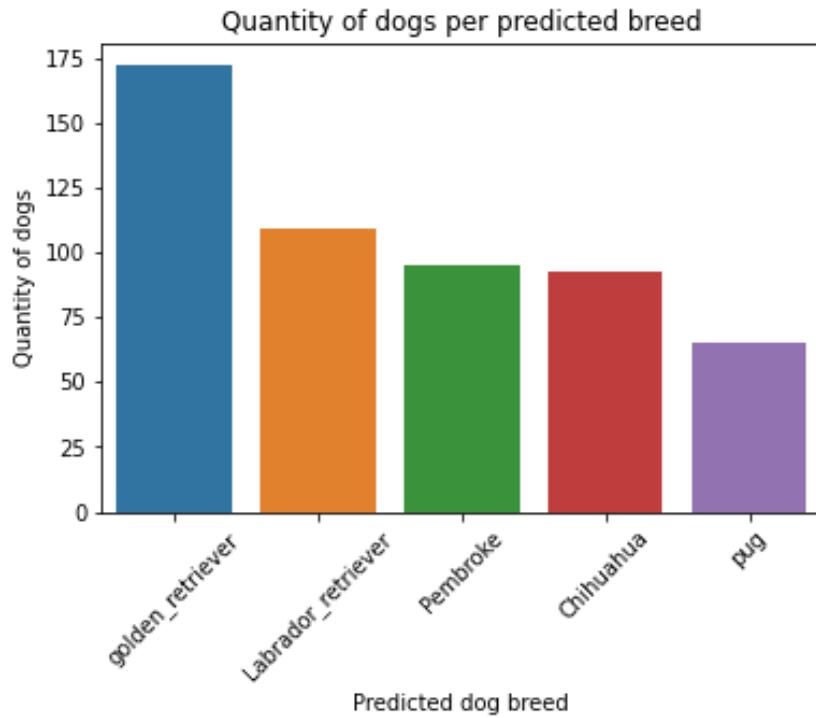
By the own definition of @WeRateDogs, from youngest to oldest are puppo, pupper and doggo, and fluffer would normally be a generic name, usually for dogs with excess fur. Common sense would say that “puppos” are more lovely, then should receive, on average, higher ratings, and that is exactly what happens: the puppos are leading the rankings with 12.04, but by only 0.04 points more than the second category, “multiple dogs” with 12. Following them are the doggos (11.85) and floofers (11.8), and the last one are the puppers, with rating 10.73 on average.



*Example of a picture with dogs in multiple stages: a doggo and a pupper.*

### 3. Predicted Breeds

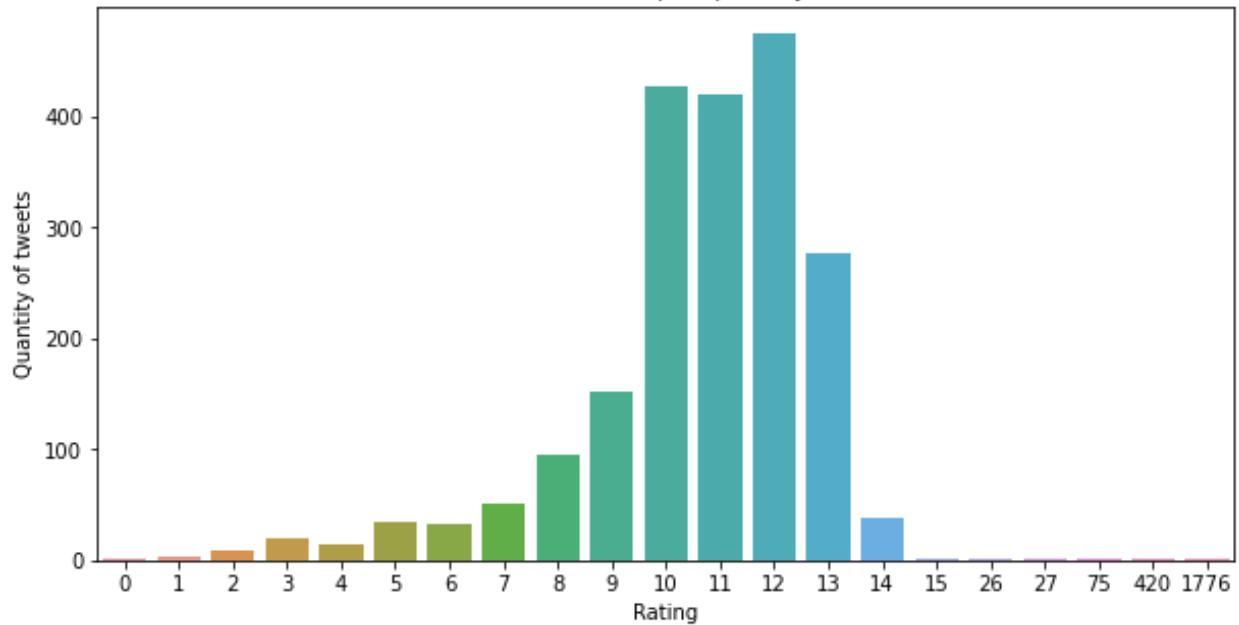
The original tweets do not have the breed of dogs posted. To solve this, we used an artificial intelligence algorithm to predict the dog breed of each picture in the dataset. Using this prediction, we can see one interesting fact: from the whole dataset of 2057 tweets, we have 1735 different dog breeds predicted, but the top 5 breeds alone represent 25% of the total. The breeds are Golden Retriever, Labrador Retriever, Pembroke, Chihuahua and Pug.



#### 4. Distribution of Ratings

If we plot a visualization of the number of tweets per rating, we can see that it is a left-skewed distribution, with first-quartile at 10, median at 11 and third quartile at 12. It clearly shows that most of the dogs are concentrated in these three ratings, having a gradual decrease in the quantity when reducing the rating, and an abrupt decrease in quantity when going for ratings higher than 13.

Number of tweets per quantity of likes



```
count      2057.000000
mean       11.696159
std        40.039408
min        0.000000
25%       10.000000
50%       11.000000
75%       12.000000
max       1776.000000
Name: normalized_numerator, dtype: float64
```