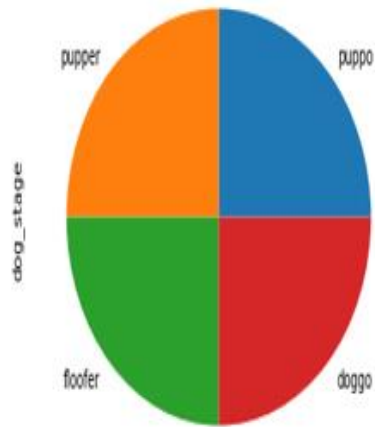# WeRateDogs Visualizations


It's clear that for the dog stages in our dataframe , we have approximately equal proportions for the different dog stages ('pupper' , 'puppo' , 'floofer' , 'doggo')

- In our data we don't know which dog stage is the most common stage that the dogs are present in , so after cleaning data it's clear that all the dogs are distributed equally through the four different dog stages.
- Each dog stage share a quarter of a pie.
- This gives as an insight that dogs have a life cycle through different dog stages.
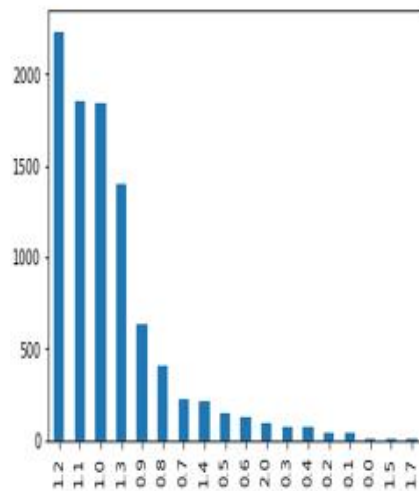
From our analysis , we can see that most dogs are rated with 12 . Also for the extreme values 20 and 0.1 it comes with a very low percentage compared to the values of 12

- In order to evaluate the dog ratings , I used a simple equation that divides the numerator over the denominator (which is 10 in our case) to represent dogs ratings.
- We can see that 12 rating is the most common repetitive rating among all the other ratings.
- Also the extremes are not too common , the very large ratings and the very small ratings are not too common in our data.
- Our ratings can represent a gaussian curve if they are drawn in that way , with a mean of 12.
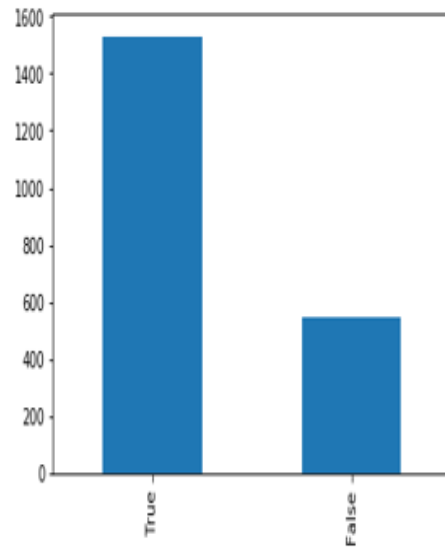
Here we can guess that the machine learning algorithim needs to be trained much more to give more accurate results or that some pictures didn't include dogs at all

- This analysis represents the number of real dogs and anything else, also it can represent the accuracy of the used machine learning algorithm in detecting dogs.
- In our analysis we have about 1500 dogs and about 600 anything else (except dogs).