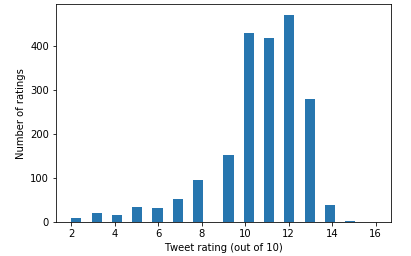
**WeRateDogs Twitter Data Insight and Analysis Documentation**

By: Jason Kim

Analysis was possible through the use of a cleaned master twitter archive file that was the result of a three-file combined merge.

Applying a statistical summary on the master dataset, we can quickly observe a few things. First, we can see that the average rating of all dogs is about 12. That seems pretty consistent with the histogram shown below. As we can see from the histogram, a rating of 12 was given the most often. Furthermore, about 90% of all ratings lie between a rating of 8 to 13. There were a few outliers that lie much further out than this visual shows, but those were omitted for the sake of analysis.



Based on the neural network’s first prediction, out of all the dog breeds between a rating of 8 and 13, the golden retriever was the breed with the greatest amount of ratings. This was followed by the Labrador retriever and the Pembroke. Furthermore, the Saluki had the most favorites on average, followed by the French bulldog and the black and tan coonhound. The standard poodle had the most average retweets, followed by the Afghan hound and the English springer.

Lastly, from the bar graph below, we can see that there is a strong positive correlation between ratings and number of favorites the dog received. The higher the ratings, the higher the average favorite tweet count. With the exception of rating 8 and 9, which are very close, we can see a clear rise with the biggest jump coming from rating 12 to 13. It was observed that a one rating increase resulted in a doubling of the average favorite count by about 10k.

