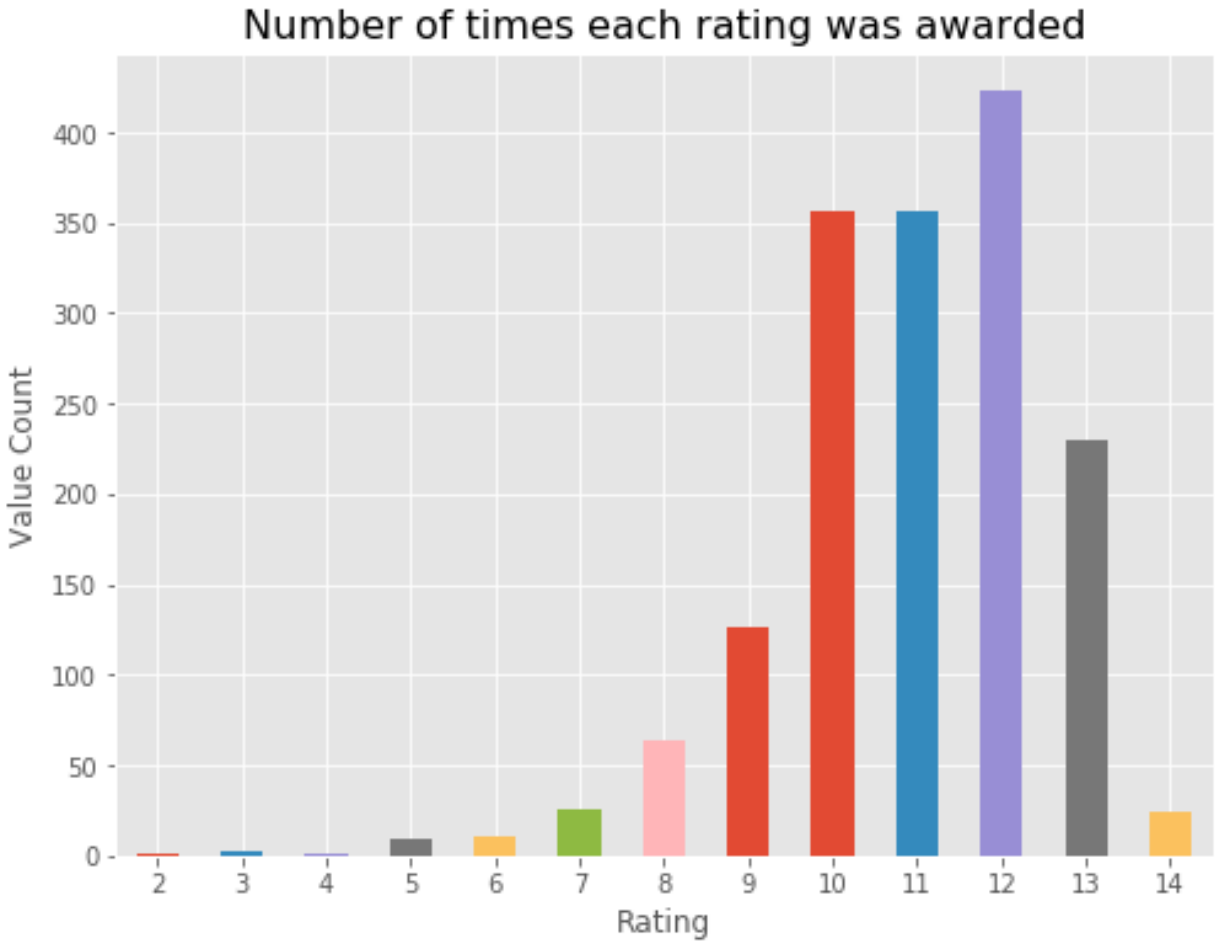
WeRateDogs Analysis & Visualization

Introduction

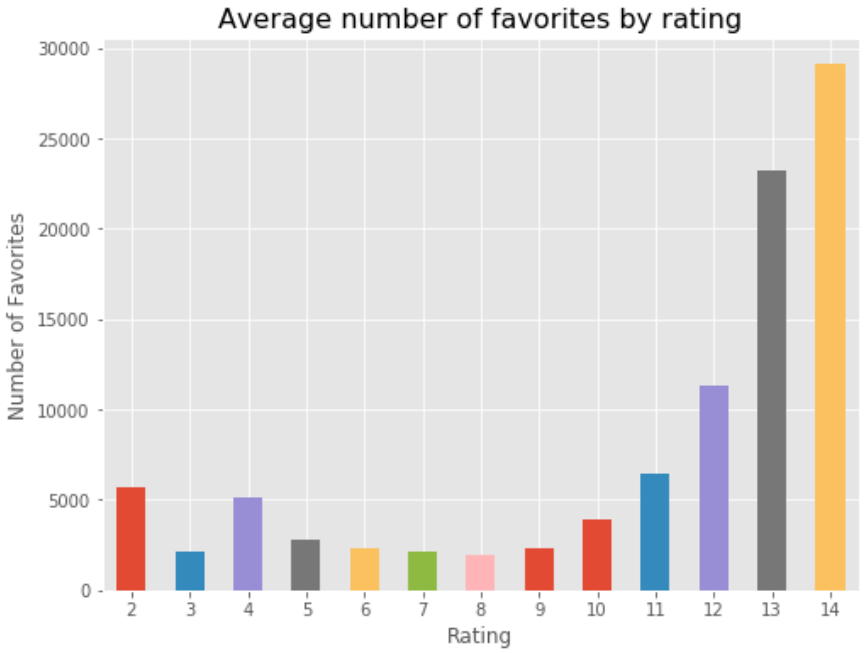
WeRateDogs is a dog rating twitter account that's gained significant popularity over the years. The twitter archive provided for analysis begins with the very first tweet which was on November 15th of 2015 and goes through August 1st of 2017. The vast majority of dog ratings fell between 2 & 14 points out of 10, so that was the range of values I used for analysis. The images from these tweets were run through Udacity's neural network to classify each dog by breed. Even image predictions with high confidence levels were oftentimes wrong, so they weren't used in my analysis.

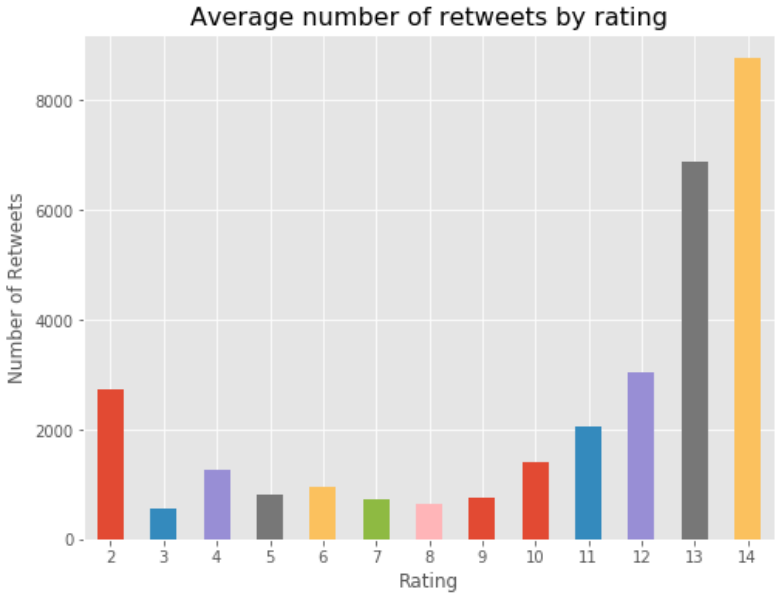
Analysis & Visualizations



While ratings were given on a 10 point scale, over 60% of dogs were given an 11 or higher. Their loose adherence to the 10 point scale, comical commentary, and cute dog pictures are what makes WeRateDogs such a popular twitter account. The distribution is skewed to the left with an average of 10.95 and a median of 11.

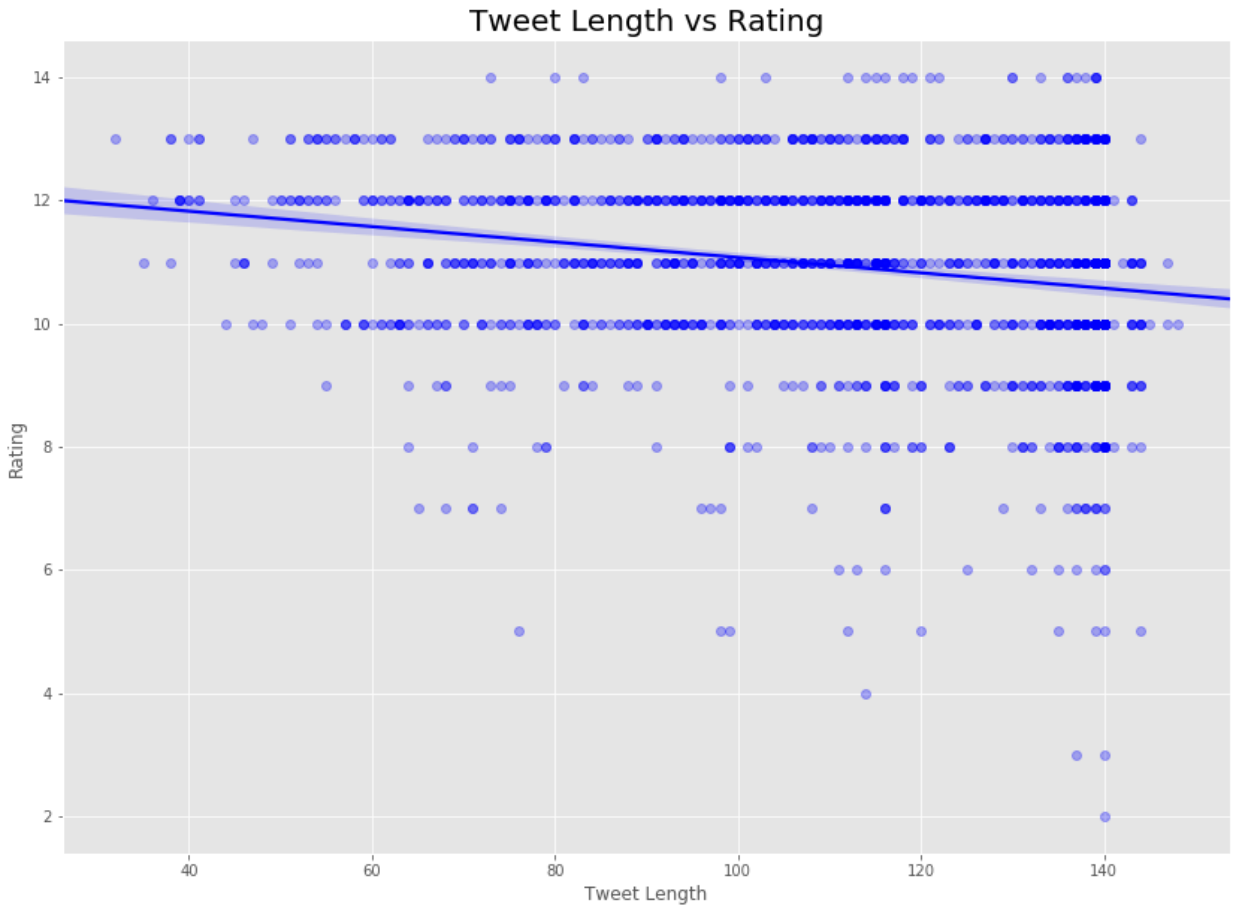
Unsurprisingly, the highest rated dogs also received the most favorites. However, from this chart it appears that low rated dogs (2's & 4's) are more popular than those awarded a 5-10 rating. Taking into account the low favorite count for 3's and the small sample size for 2's & 4's (1 record each), it seems likely that these average values would decrease with more data.



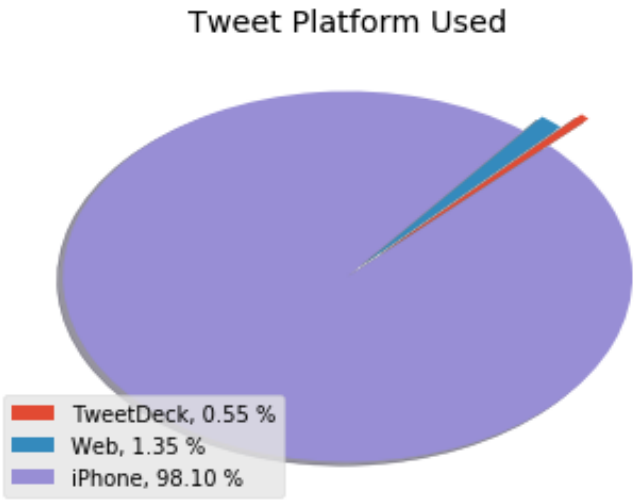


Only a 2 for this puppo?

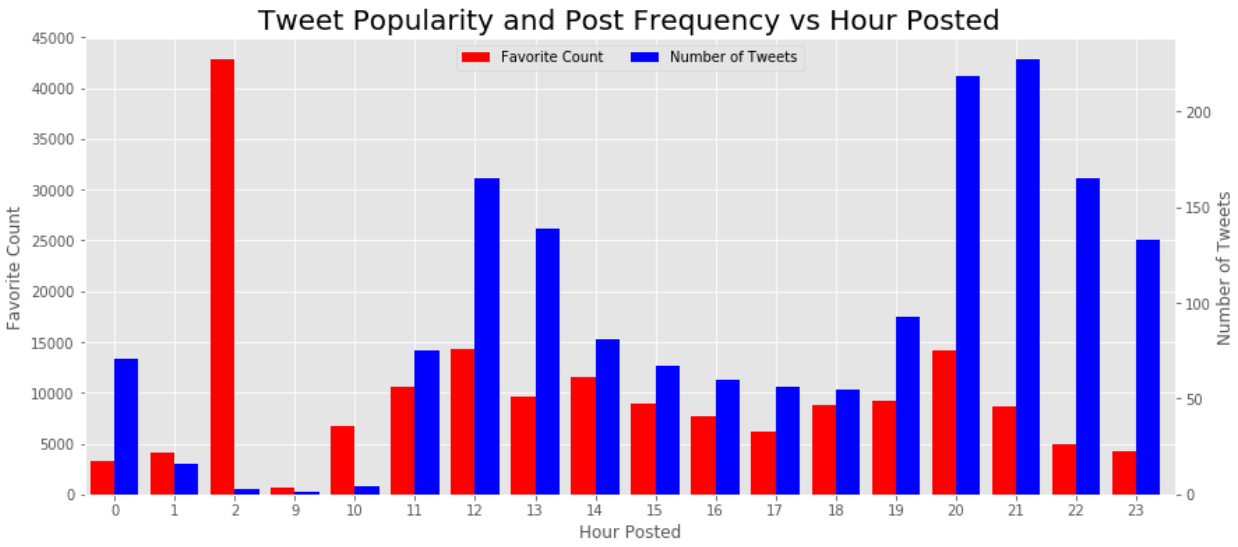
The graph for the average number of retweets looks very similar to the one for favorites, except the popularity of the tweet with a 2 rating was even more pronounced on this graph. When we look at the tweet, it's easy to see why it was retweeted so many times. The dog was actually pretty cute, but may have been acting a tad inappropriate in the picture.



There's a weak negative correlation (-0.21) seen between tweet length & rating. This makes sense when you think about the explanations behind the ratings. WeRateDogs could post a cute dog with a high rating and people would get it, but their followers would want an explanation for a low rating.



The overwhelming majority of tweets were posted with an iPhone, with only a few coming from the web & TweetDeck. TweetDeck is a customizable dashboard hosted on twitter.com.



Although tweets posted at 2AM had the highest average, there were only 3 of them and one was a serious outlier. Had the one with 121,359 favorites been omitted, the average would only be 3,547 instead of 42,818.

Conclusion

WeRateDogs is a highly successful twitter account that's managed to amass a huge following almost entirely from their iPhone. They're clearly doing things right, but I think they could do even better with a little help from data analysis.

If it were my twitter account, I'd be a little more liberal with the 13's & 14's. The dogs that received 10's were just as cute, if not cuter than the 13's & 14's, but they were favorited & retweeted significantly less. As an authority on cute dogs, people will favorite & retweet dogs that WeRateDogs thinks are cutest. Not needing a lengthy explanation for lower ratings is just an added bonus. Of course they couldn't do this will all dogs without diminishing the value of 13's & 14's, but a small increase in high ratings should be fine. To maximize engagement, I would save the cutest dog pictures and schedule them for 12 PM & 8 PM when the average favorite count is highest.