WeRateDogs Twitter Archive Analysis
By Aastha Arora

@dog_rates
WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog.
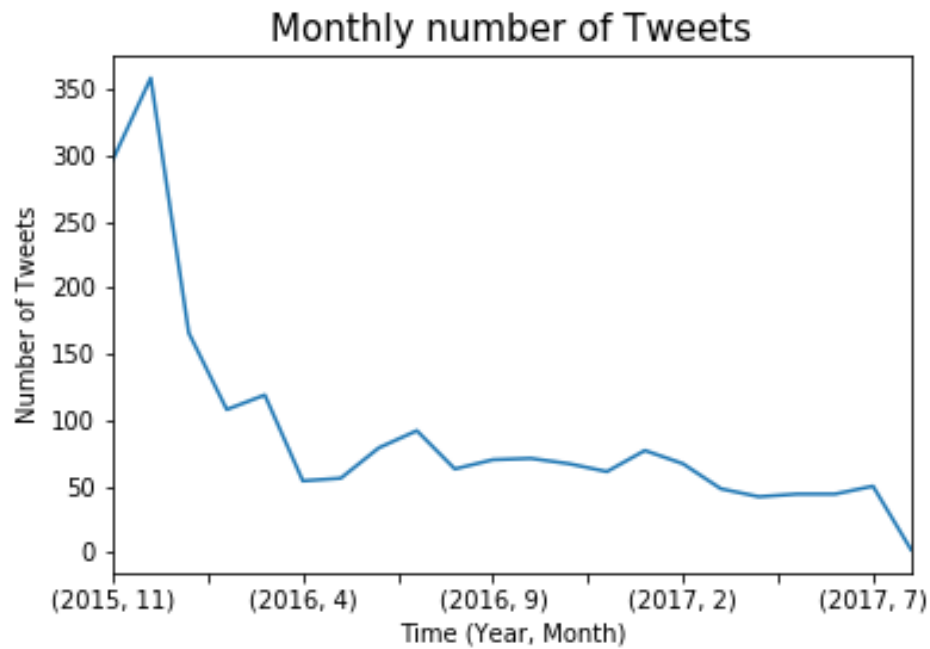
The tweets have a unique rating system. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc.

Twitter archive data for all the tweets by WeRateDogs till August 1, 2017 was collected and analyzed. The data was filtered to include only the original tweets (no retweets) that have images. This data was merged with the image prediction data provided by Udacity.
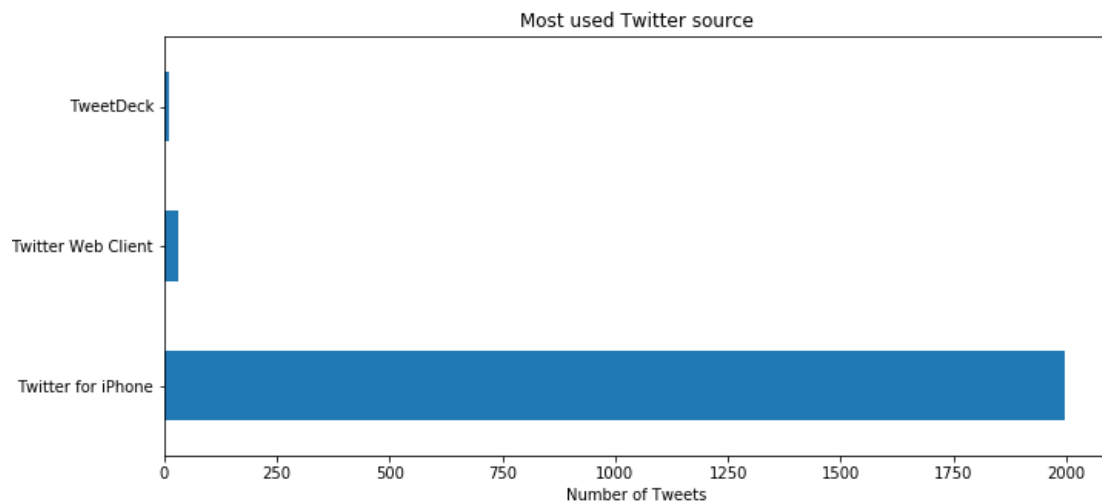

Sample Tweet

The number of tweets published by WeRateDogs decreased over time. The line charts shows the trend in the number of tweets over the different months.
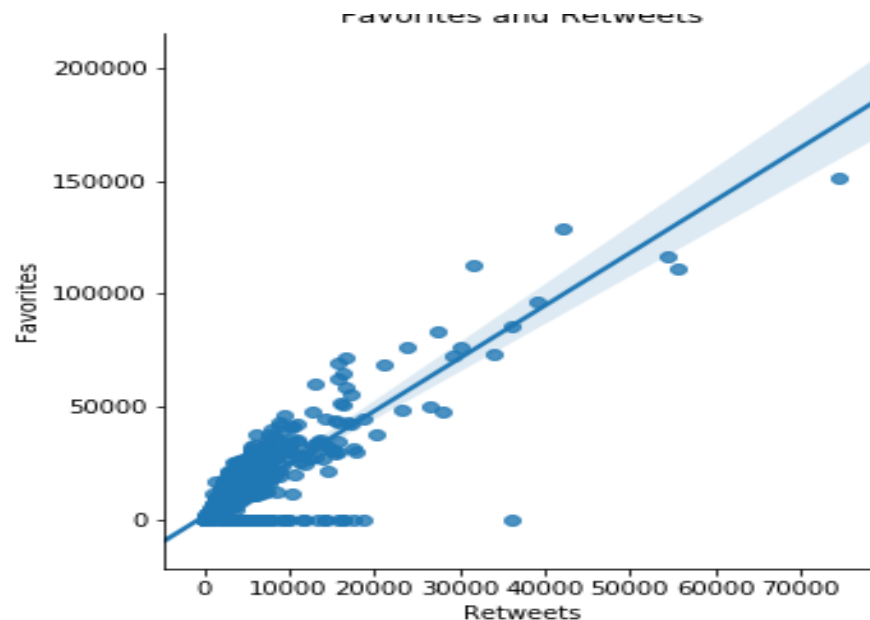

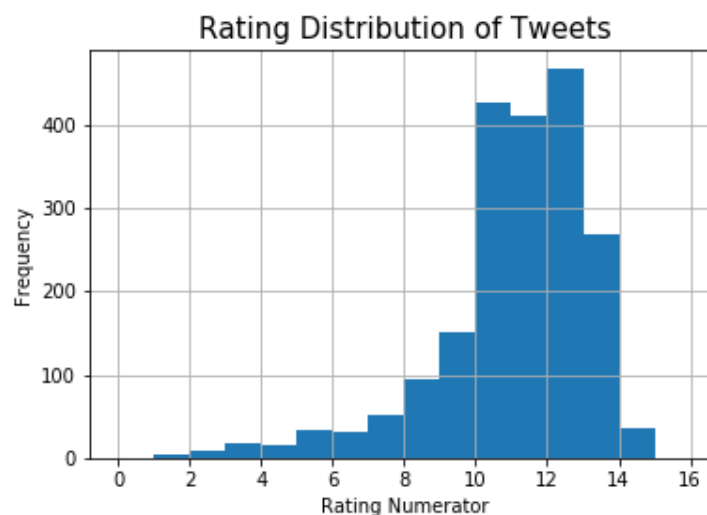Data was analyzed to identify the source of the tweet device.

Most of the tweets published by WeRateDogs were using Twitter App for iPhone.

I wanted to see if there a relationship between the number of retweets and likes received on a tweet. To check that, I plotted a scatter plot between the number of retweets and favorite count (likes).

Favorites and Retweets

As expected, there is a positive correlation between the retweets and favorites. The correlation was calculated using the corr function and it came out to be 0.86. This verifies the strong position correlation between the two variables.

Now, let's take a look at the unique ratings given by WeRateDogs. Two of the rating were outliers and were removed to visualize the rating distribution using a histogram. The most common rating given by WeRateDogs is 12. Rating 10 is close second while rating 11 is the third most common rating.

Rating Distribution of Tweets

The data collected by the Twitter API included the image URLs for the images published in the post.

Check out some of the prediction made by the results of the neural network algorithm provided by Udacity.

tweet_id	tweet_url	media_url	first_prediction	first_prediction_confidence
890006608113172480	https://t.co/dVPW0B0Mme		Samoyed	0.957979
889665388333682689	https://t.co/BxvuXk0UCm		Pembroke	0.966327
889638837579907072	https://t.co/f8dEDcrKSR		French Bulldog	0.991650
889531135344209921	https://t.co/y70o6h3isq		Golden Retriever	0.953442
888078434458587136	https://t.co/DK7iDPfuRX		French Bulldog	0.995026

Lastly, I was curious to check out the most retweeted tweet by WeRateDogs. It turns out that it is actually a video of dog swimming in a pool. In fact, even the second most retweet post is a video of a dog making bubbles in water. Check out the URLs for these tweets at https://t.co/7wE9LTEXC4 and https://t.co/YcXgHfp1EC

Complete dataset and code is available in the wrangle_act jupyter notebook.