ACT REPORT (PROJECT #2)
## We Rate Dogs Twitter Data Analysis

Overview

WeRateDogs is a Twitter account that rates users' dogs and adds a lighthearted comments. The denominator of these ratings is almost always 10. however, the numerators? frequently greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs, Brent." WeRateDogs has over 4 million followers and has received international media coverage.

For Project 2 in the Udacity Nano-degree I analysed the Twitter data from the account to master data gathering and wrangling skills. Furthermore, the analysis of this twitter data helped to generate insights and answer questions like:

- What is the most common dog stage in the tweets?
- What type of dog breed gets the most likes?
- Is there a correlation between the number of retweets and favourites count?
- Does the Source of the Tweets influence popularity? (retweet count)
- What is the most common value for ratings

Most common dog stages:

I analysed the data to find the most common dog stages; and visualized the result in the plot below:
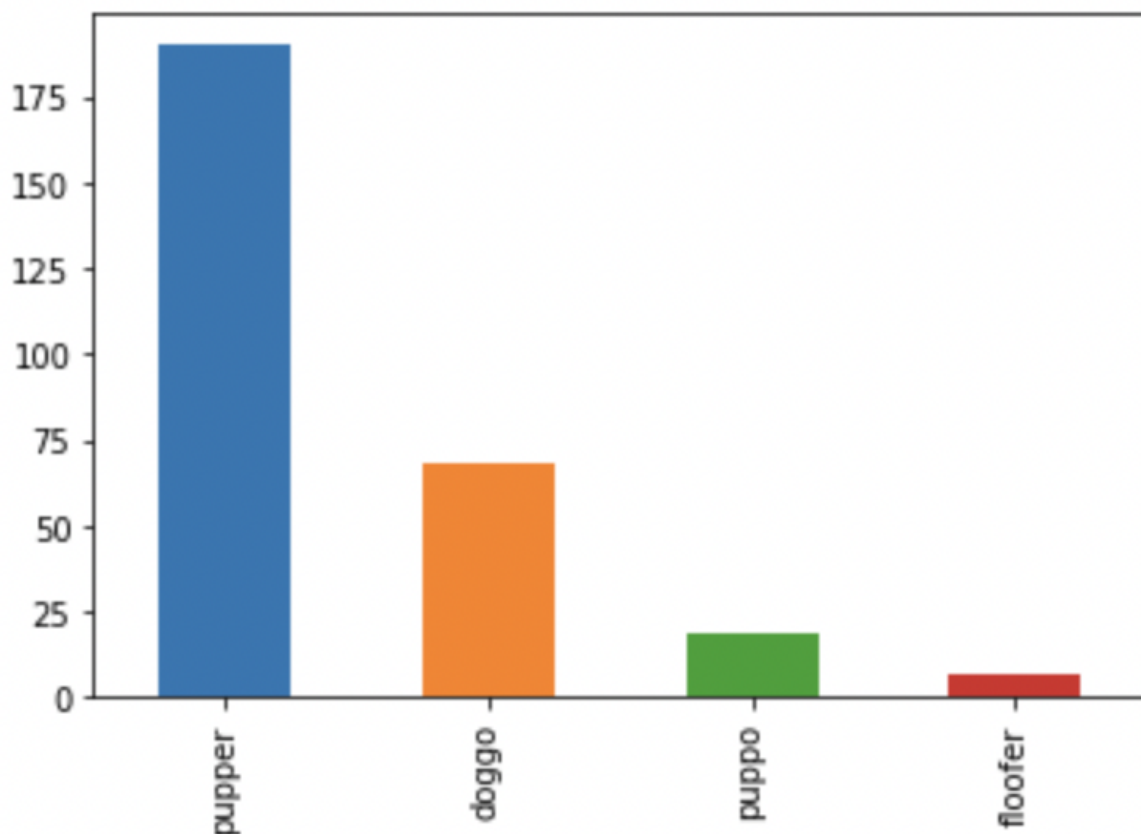

The pupper dog stage appeared to be the most popular stage among the tweets.

Correlation between the number of retweets and favourites count?

In my Analysis, I was curious to find out the relationship between the number of retweets and favourites. Therefore; I plotted a correlation graph to explore this relationship. Below is the plot:

The Correlation Between favorite counts And retweet counts is 0.925835521579
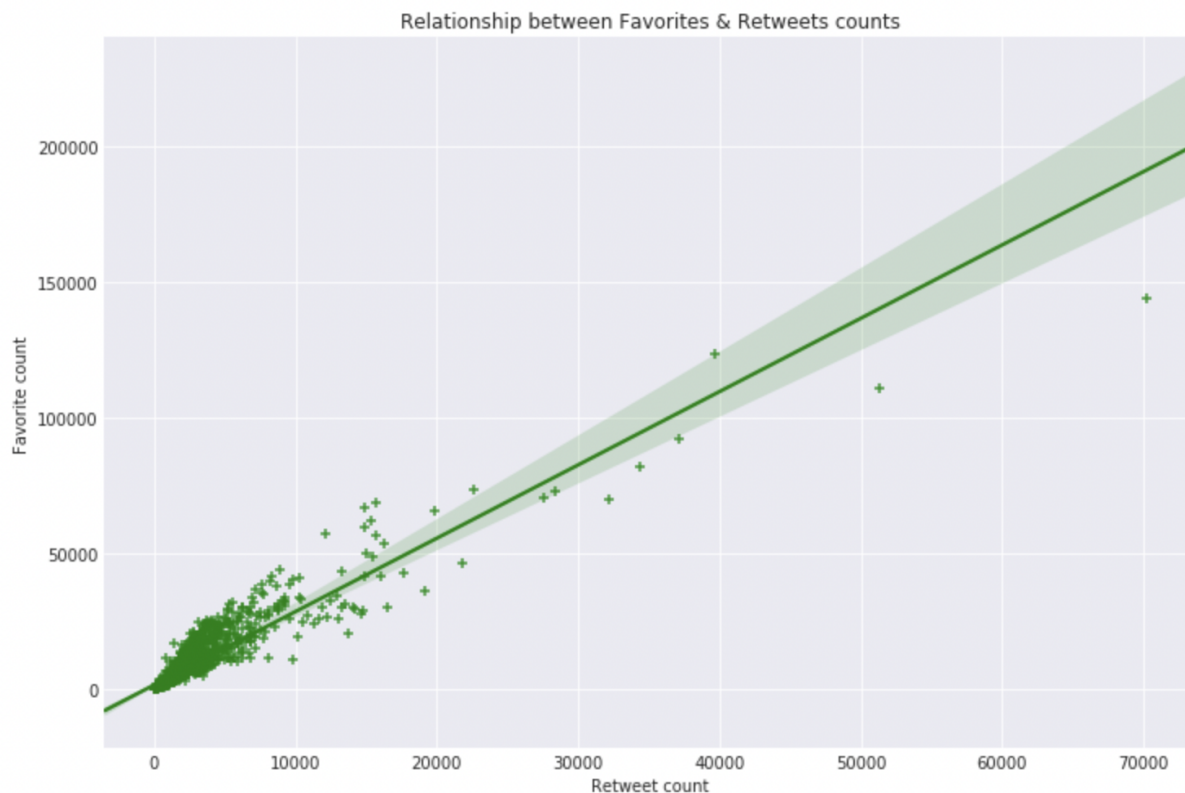
The regression was done using seaborn module and the regplot function. From the results, we observe a high correlation relationship between the number of retweets and favourites.

Therefore; it is safe to assume that Twitter users that like a post are more likely to repost it on their timeline to show to others.

Does the Source of the Tweets influence popularity? (retweet count)

Another interesting insight gathered during the analysis was the answer to the question of the virality of tweets. So I sought to find out if the source of the tweets from different

devices had anything to do with retweet counts. The retweet count was my chosen indicator for virality. The results are displayed below:
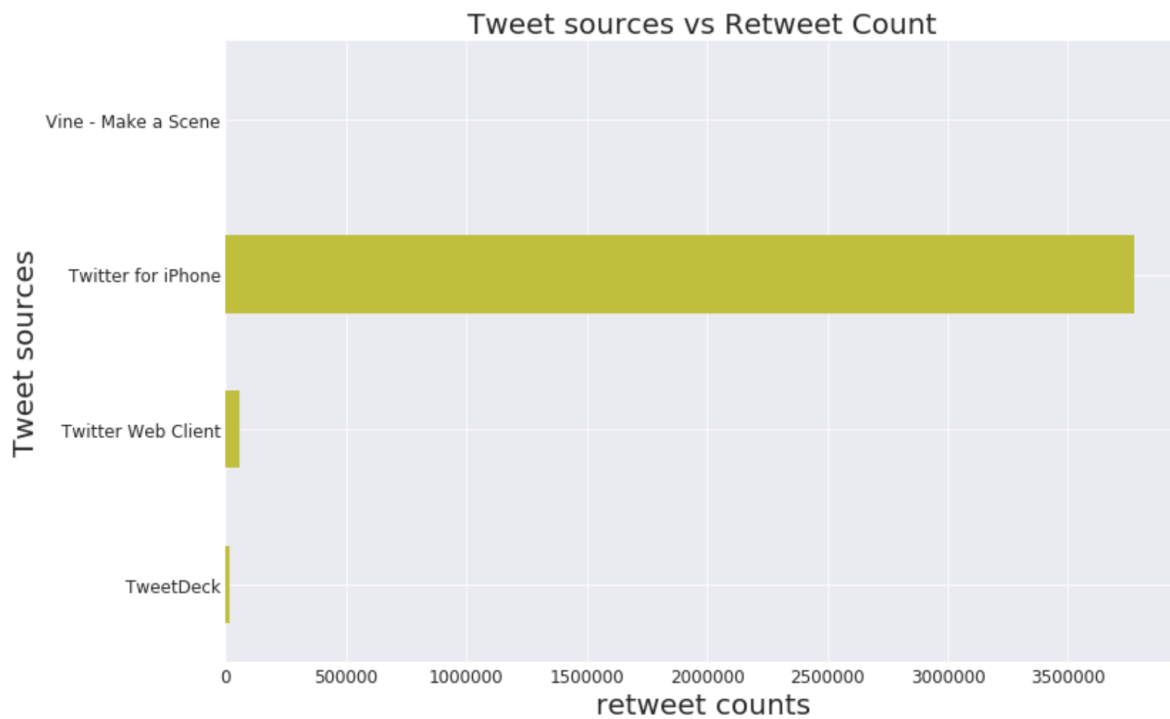
The results show that tweets made from the iPhone had the most retweet counts. Therefore; we might conclude that to get a viral tweet; iPhones might have an influence on the reach.

What is the most common value for ratings

Finally, I explored the dataset for the most common value for ratings; From the visualization plot below, the common rating scores fell in the range between 9 and 13.

Ratings frequency
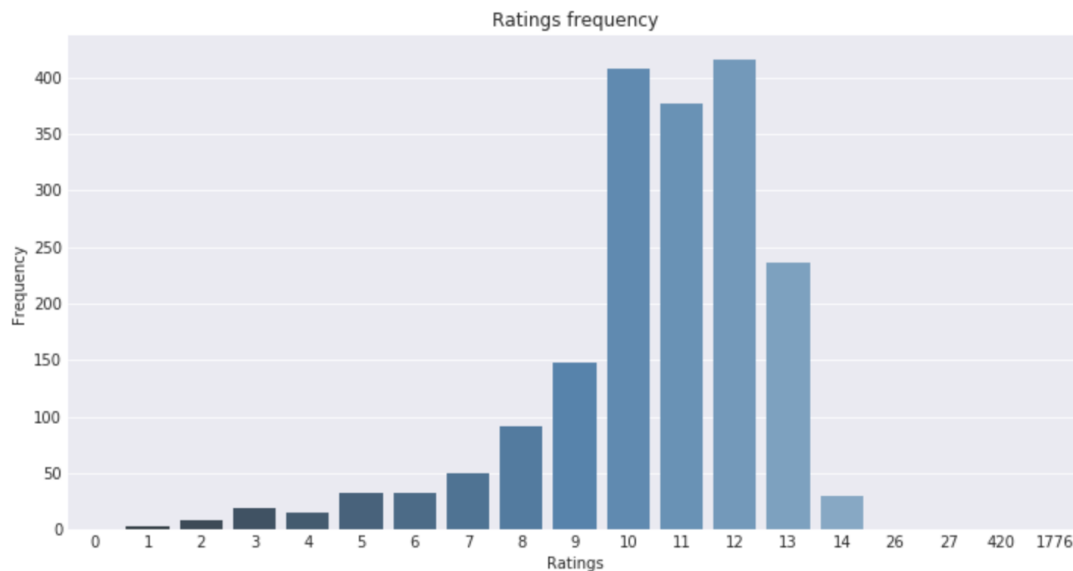
Conclusion

In Summary, this analysis was really enjoyable. In fact, analyzing such fascinating data sets encourages me to pursue a career in data analysis. I wish WeRateDogs continued success and that those who were previously unaware of its efforts would now be more interested. Visit them on Twitter and enjoy using the tweet and retweet functions!