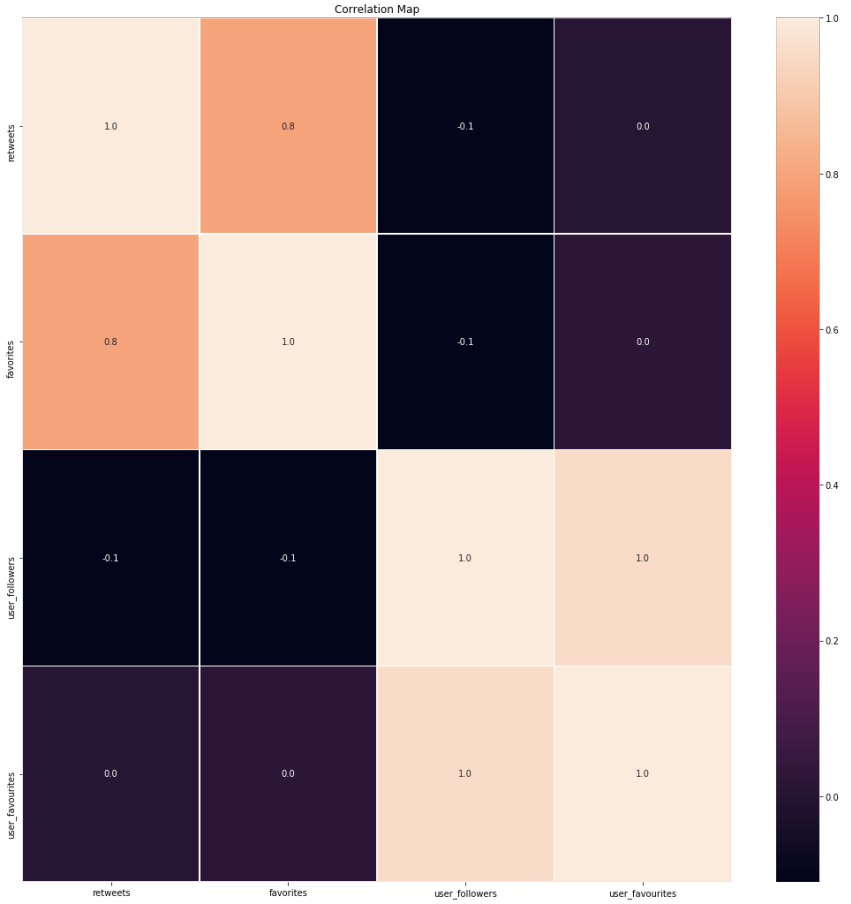
Analysis and Visualization of WeRate Dogs Dataset

Introduction:

The dataset we analyzed (and analyzing and visualizing) is the tweet archive of Twitter user **@dog_rates**, also known as **WeRateDogs**. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always 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.
WeRateDogs **downloaded their Twitter archive** and sent it to Udacity via email exclusively for you to use in this project. This archive contains basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017. More on this soon.

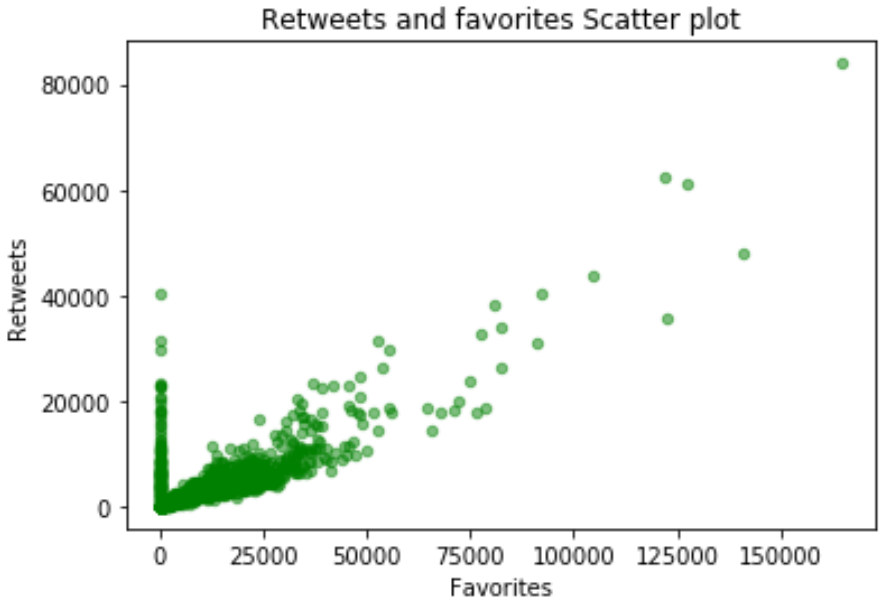
Analysis and Visualization:

In this project, retweets, rating_numerator, favorites, user_followers and user_favourites were used for generating heat map



The above heat map indicates that retweets and favorites has strong positive correlation (r = 0.8). As tweet gains favorites, one can expect retweets to increase and vice versa. Ratings did not get affected with any other variables.

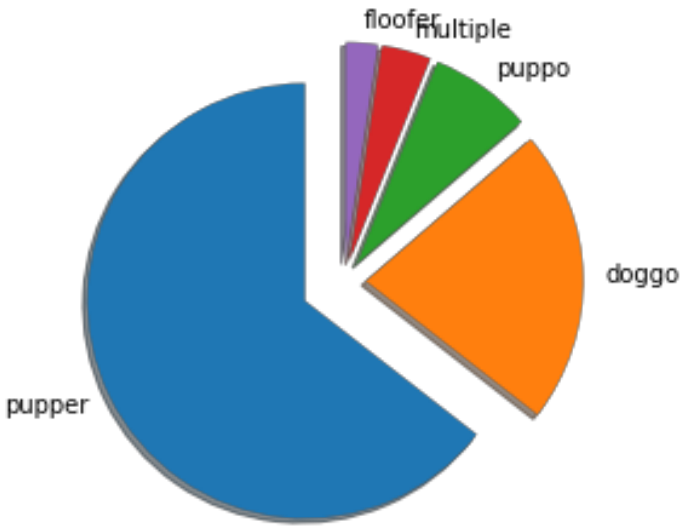
Below figure also indicates that the more favorites means more retweets.

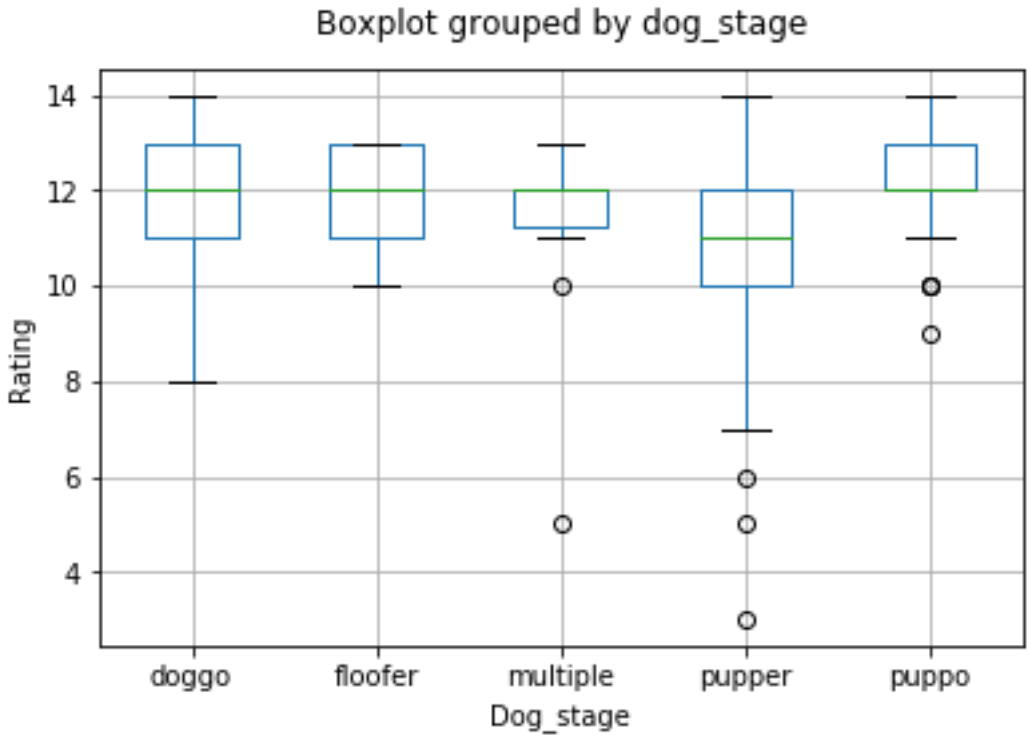


Below figure shows rating over the period of time. Here ratings upto 16 were considered for analysis. More than 75% of the data has more than 12/10 rating.

Dog Stages:

The dog will express different needs during their life stages from baby pup to elderly matriarch. Below pi chart shows that the pupper represents most part of number and interestingly it has low rating also.


floofer
multiple
puppo
doggo
pupper


Boxplot grouped by dog_stage
Rating
14
12
10
8
6
4
doggo
floofer
multiple
pupper
puppo
Dog_stage