Act Report

The Wrangle Act project consisted of gathering and loading in farious data about the WeRateDogs Twitter account. From there on, the data was cleaned and organized into a format that would facilitate data visualization and exploration.

After calling describe on the pandas dataframe I created to contain all the pertinent fields, I noticed that one of the tweets was favored over one hundred and twenty thousand times. I decided to use that favorite_count to pull up it's tweet_id, and from there the url of the image that was included in the tweet to see what kind of a dog it was of.

I also noticed that this particular tweet also had the highest number of retweets of any other in the sample. This observation proded me to look further into a potential correlation between the two parameters. It turned out that there is a very high correlation coefficient between retweets for a particular tweet and the number of times it has been favorited. In this particular case it was over 0.92. This does intuitively make since since both of these actions correlate with exposure a tweet has had and also express a similar sentiment. Someone who enjoys a tweet will either like or favorite it -- and on occasion do both!

I also thought that there may be a relationship between the number of photos included in the tweet and the number of times it was favorited. I made the assumption that users who enjoy seeing cute dogs also enjoy seeing A LOT of cute dogs. However, it turned out that there wasn't a significant correlation between these two features. Perhaps it is due to the fact that there simply aren't a whole lot of tweets with multiple photos. There is also a possibility that when scrolling down their feed a follower of WeRateDogs only sees one photo at a time and the decision to favorite and/or

retweet is independent of the number of other photos included. To explore these assumptions further, I also created a seaborn pairplot.