After the wrangling phase, I decided to do some exploratory data analysis to give me some insight into the data I had spent a lot of time gathering, accessing and cleaning. I discovered the following:

Insights:

1. There a positive correlation between the retweet count and favorite count. The higher the retweet count, the higher the favorite count.

2. Wednesday has the highest retweet and favorite counts on an average, and so if WeRateDogs wants their tweets to have maximun engagements they should tweet more on wednesday.

3. The dog breed with the highest count in the dataset is the golden retriever with 150 observations, the second highest is the labrador retriever with 100 observations and the third highest is the pembroke with 89 observations.

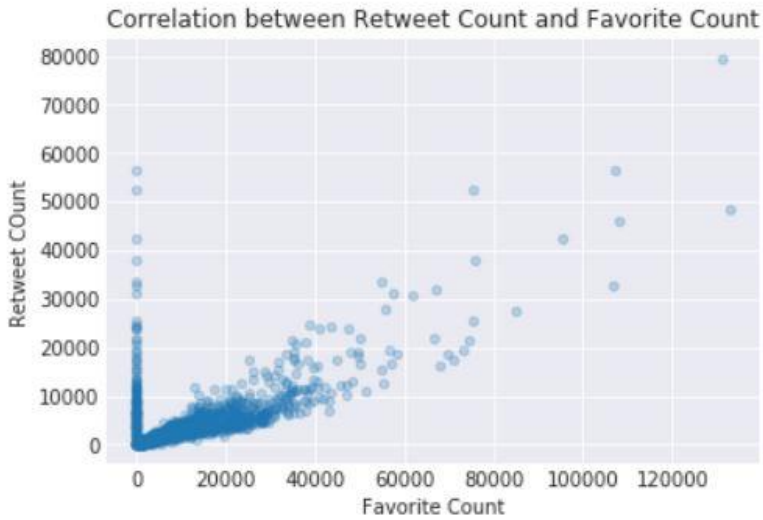
4. The dog with the highest favorite count is a lakeland terrier with tweet id 822872901745569793 and the dog with the highest retweet count is a labrador retriever with tweet id 744234799360020481.

5. The dog with the highest rating in the dataset is Atticus with a rating of 1776/10.

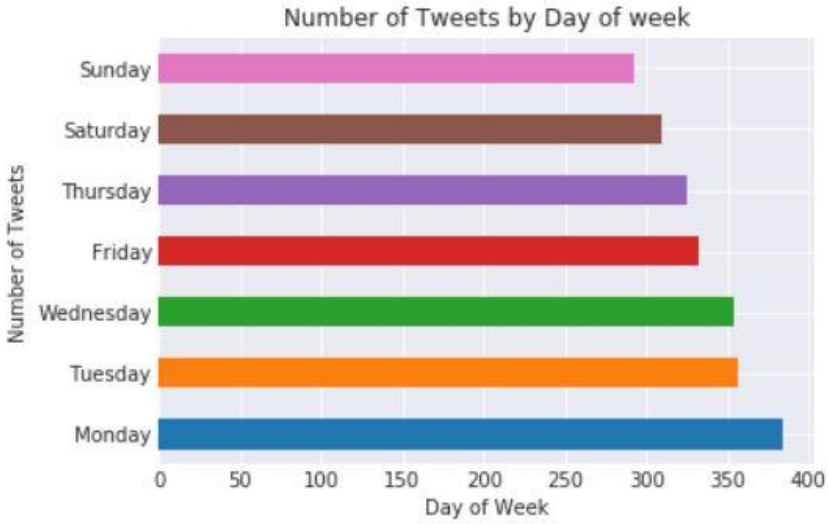
6. The number of tweets by WeRateDogs reduces as the week progresses with monday having the highest number of tweets and sunday with the lowest.

Visualizations:

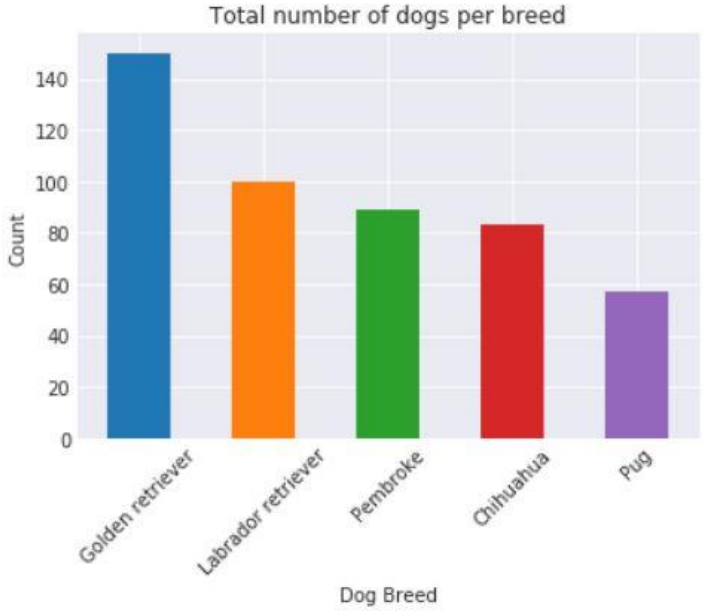
The following visualizations were derived from the available data:



The scatter plot shows a positive correlation between the retweet count and favorite count. This means that as the retweet count increases, the favorite count also increases.



This horizontal bar chart shows that WeRateDogs number of tweets gradually reduces as the week progresses, with Monday having it's highest number of tweets.



The bar chart shows the top five dogs in the dataset based on count. The dog breed that has the highest number of tweets in the dataset is the golden retriever with about 150 image predictions.