WeRateDogs Report Analysis

This is a real world data wrangling project using three different data set obtain from WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dogs. The three data set was provided by Udacity for analysis. They include the tweet_json, image-prediction and twitter-archive-enhanced data set. After cleaning the data set we are able to get the following insight from the data set
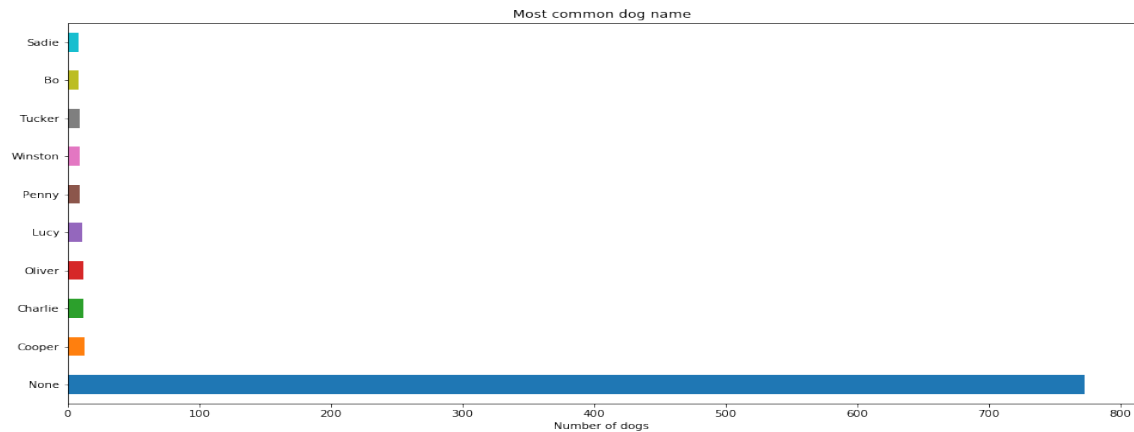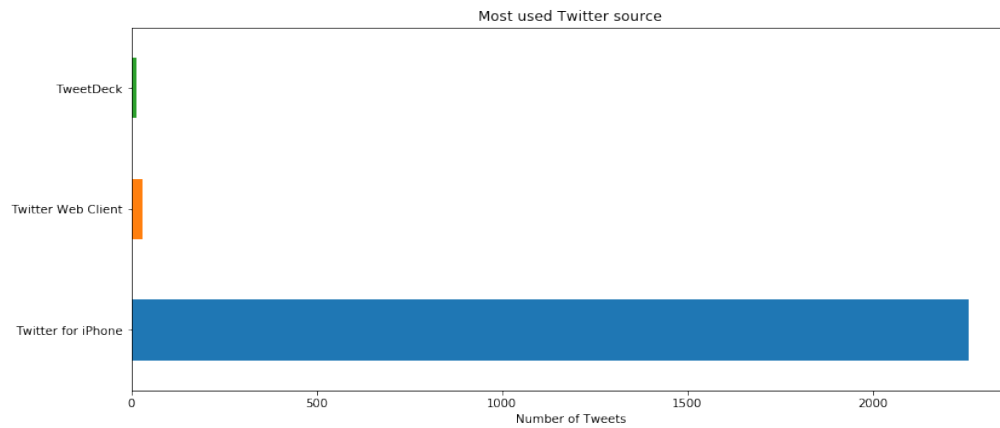

The most common dog name

The most common dog name are Cooper, Charlie, Oliver and Lucy
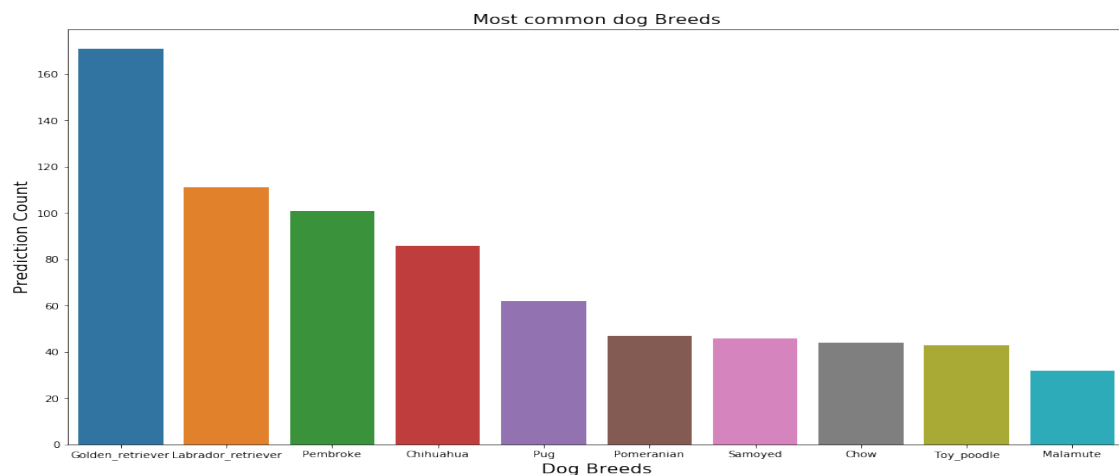
Most common dog breeds

The most popular dog breeds is the Golden_retriever.

Most used Twitter source

The most used tweeter source

The most used tweeter source is Twitter for iPhone which amount to at least 97% of the tweet posted between 2015 to 2017. This indicate that a lot of people tend to use their iPhone to react to the rating of the dogs


WeRateDogs tweeting trend over time
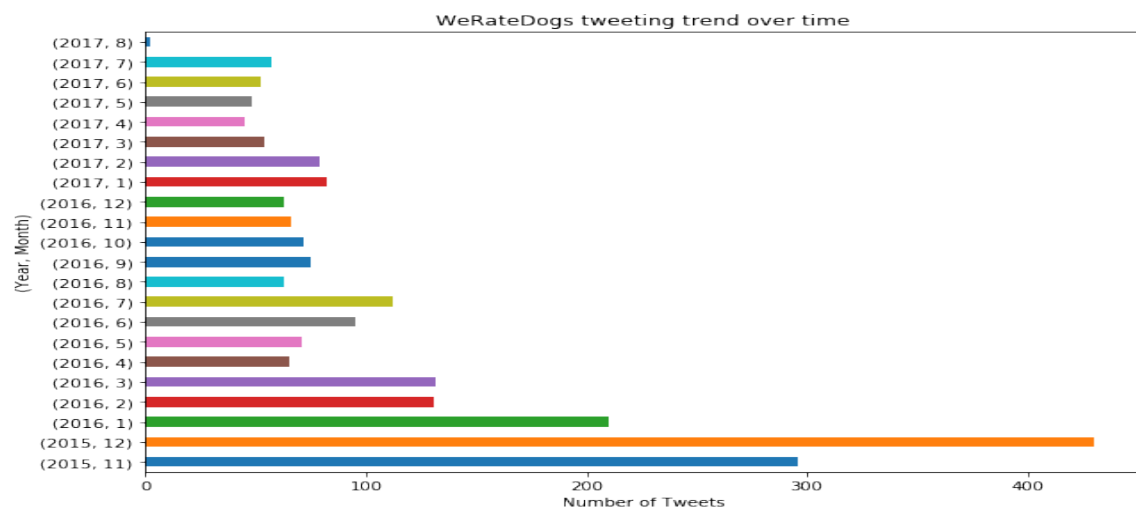
Tweeting trend over time

The highest tweeting trend was in December 2015 which as 430 tweeting trend. The closes tweeting trend was 296 which was November 2015. As we can observe that the tweeting trends decreases with time.

The Dogs Rating Distribution
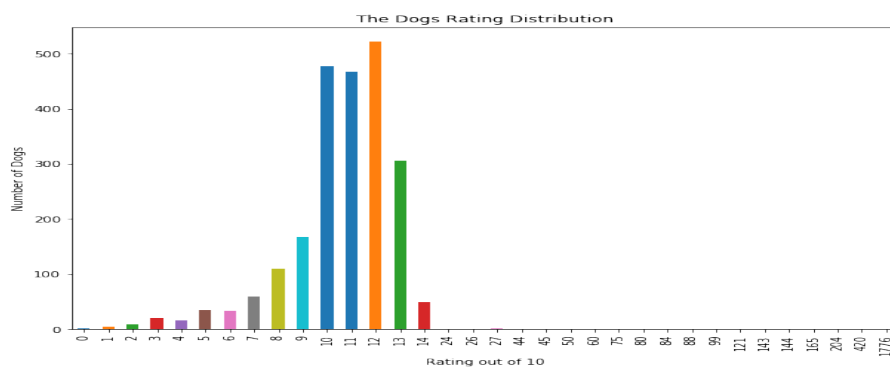
Dog rating distribution
The highest number of rating distribution is on the scale of 12.

Favorite Count vs. Retweet Count
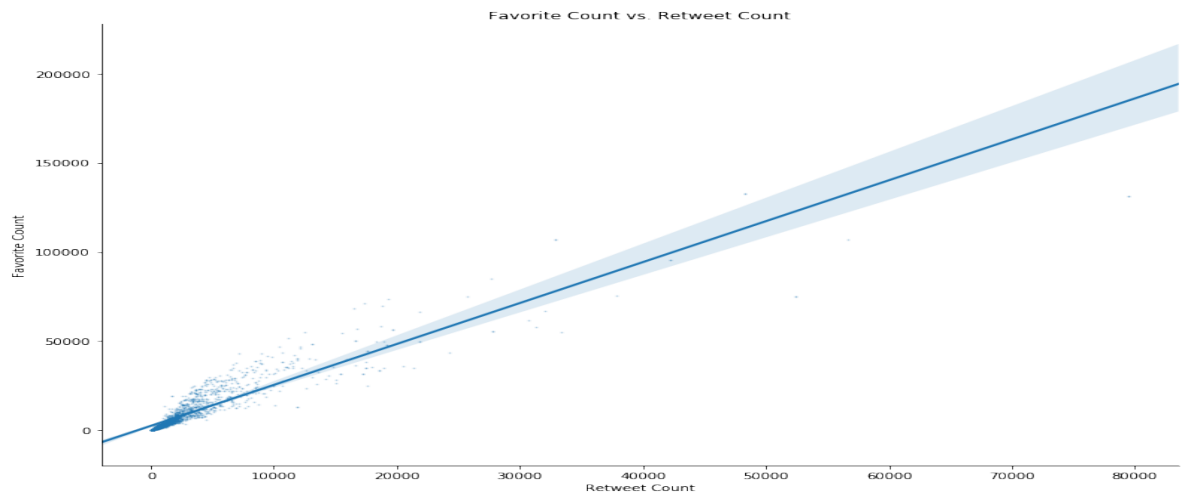
Favorite count and Re-tweet count
The retweet_count has a positive correlation with the favorite_count