

## Act Report

The objective of this report is to draw insights from the data wrangled in wrangle\_act.ipynb. The aim was to gather, assess and clean data from three different sources, all associated to the WeRateDogs twitter page, and create interesting and insightful analyses and visualizations from the cleaned data source.



**Tools used:** Python, NumPy, pandas, Matplotlib, Jupyter notebook, Excel

### Introduction:

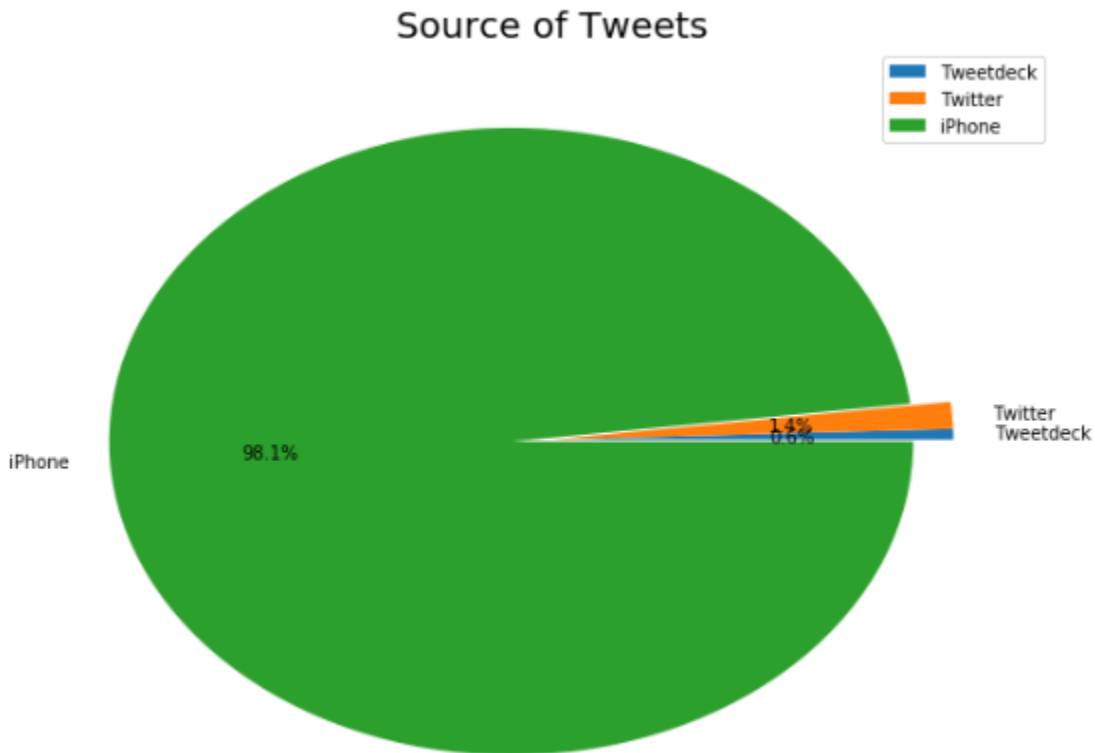
The WeRateDogs twitter archive contains 2356 tweets with ratings, and from each of the three sources we are given (tsv, csv, txt), useful data such as tweet id, tweet date, dog breed, retweet & favorite counts can all be gathered. The wrangling process involved wrangling the data, which started off by joining the three sources together as one master file, then further analysis & cleaning was carried out to make the file easier to interpret and work with.

## Insights Gathered:

### The source of the Tweets

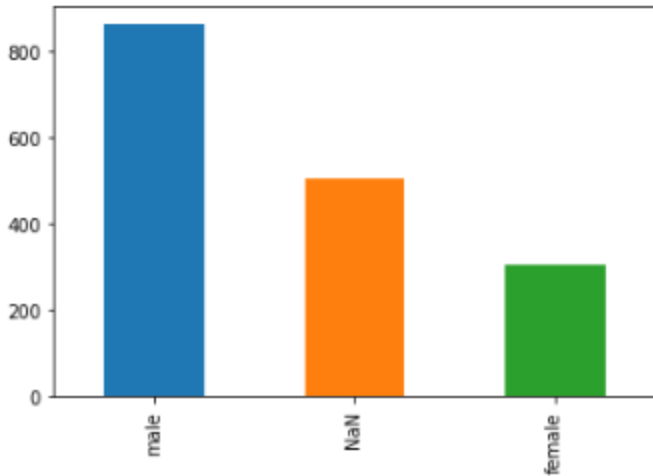
It was interesting to find that 98% of the tweets sent, came from twitter via iPhone. The remaining sources, TweetDeck and Twitter accounted for less than 2%.

```
Text(0,0.5,'')
```



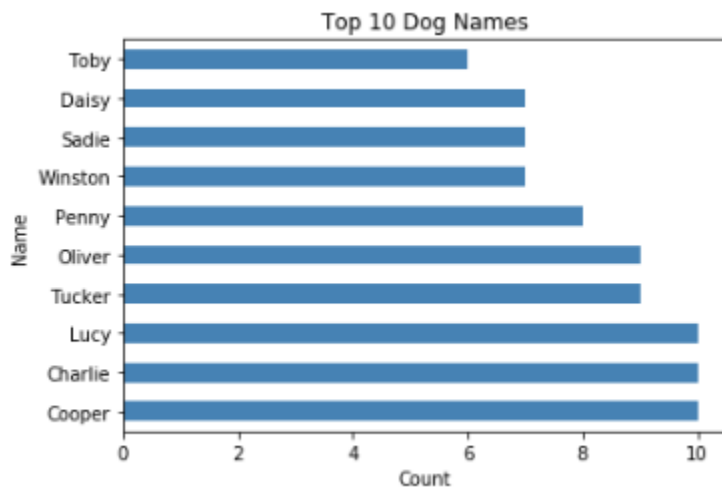
### Gender of the Dogs in the Tweets

It was found that most of the tweets sent were of Male dogs, accounting for almost 900 in total. Female's accounted for just over 200, while the remaining 500+ tweets had an unknown gender.



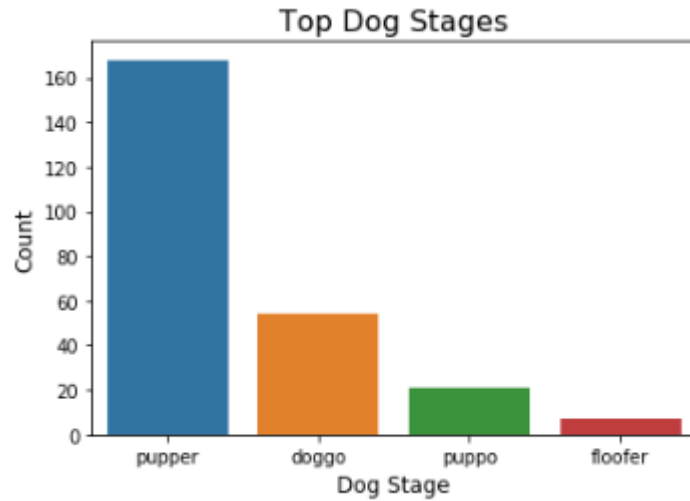
### Most Popular Dog Names

It was found that there was a four-way tie for the most common dog names from the tweets. Those names were Lucy, Tucker, Charlie and Cooper, and they were found 10 times each.



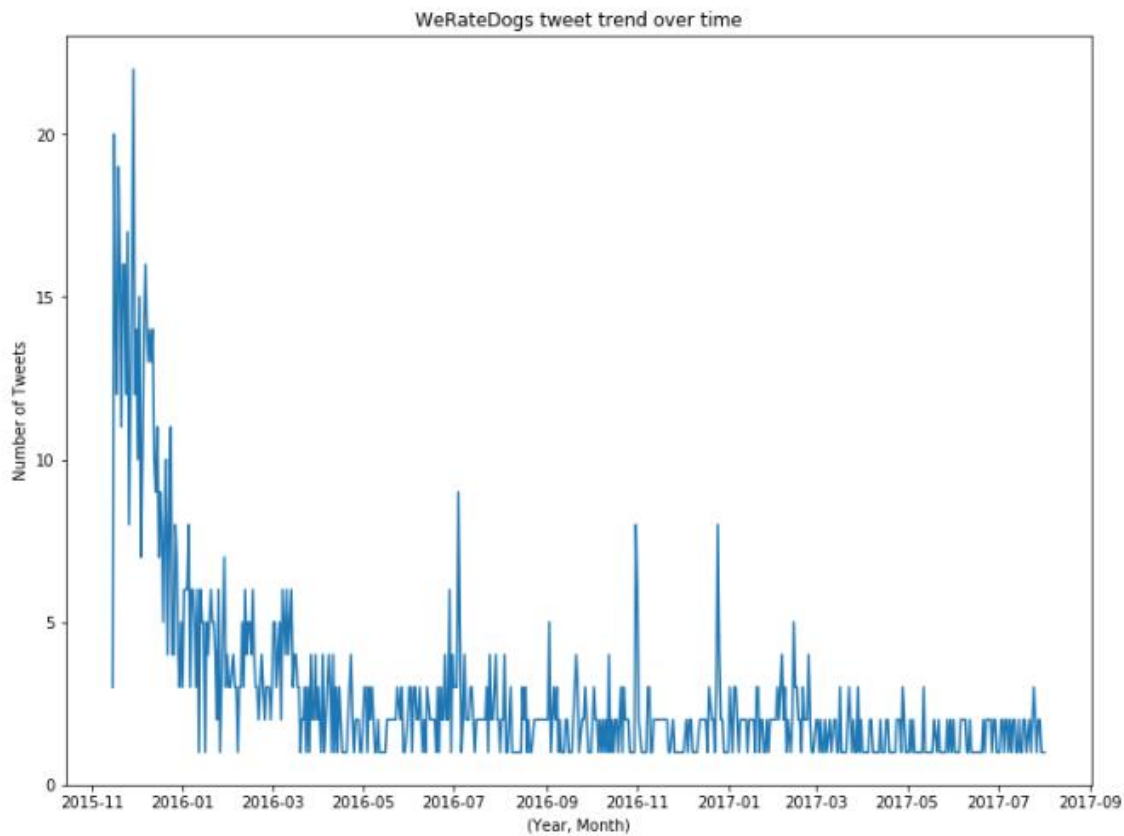
### Most common dog stage found

Dogs that belonged to the Pupper dog stage were by far the most common from the data collected, while Floofer was the least popular.



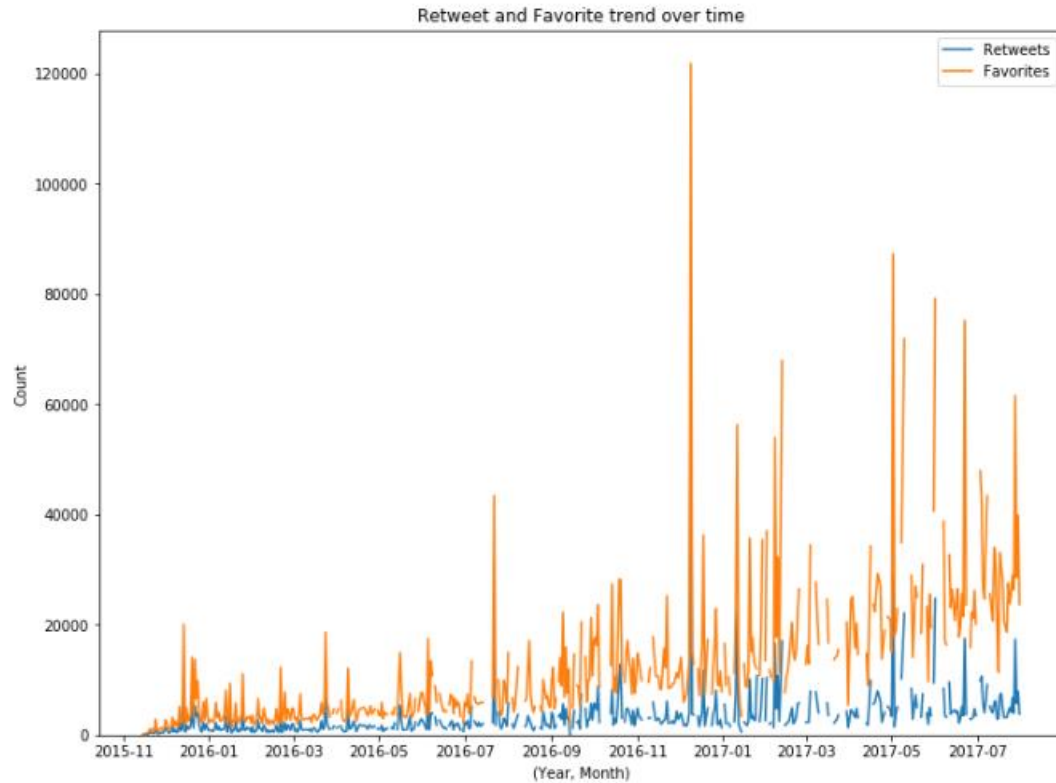
### Tweet Trend over time

There was a large spike in the number of tweets sent at the start when data was first collected in November 2015. This number fell dramatically over the next number of months. It has however levelled out coming into 2017.



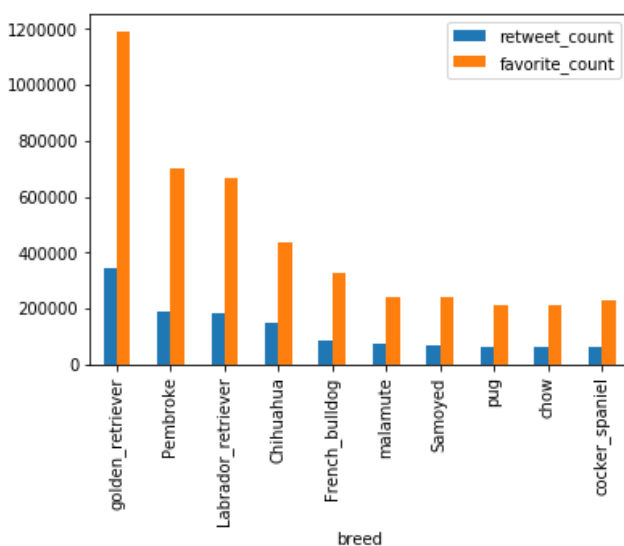
### Retweet and Favorite trend over time

Large increase is notable in the average number of favorites on tweets since the data was first collected. This proves that the twitter account is becoming more people as time has went on.



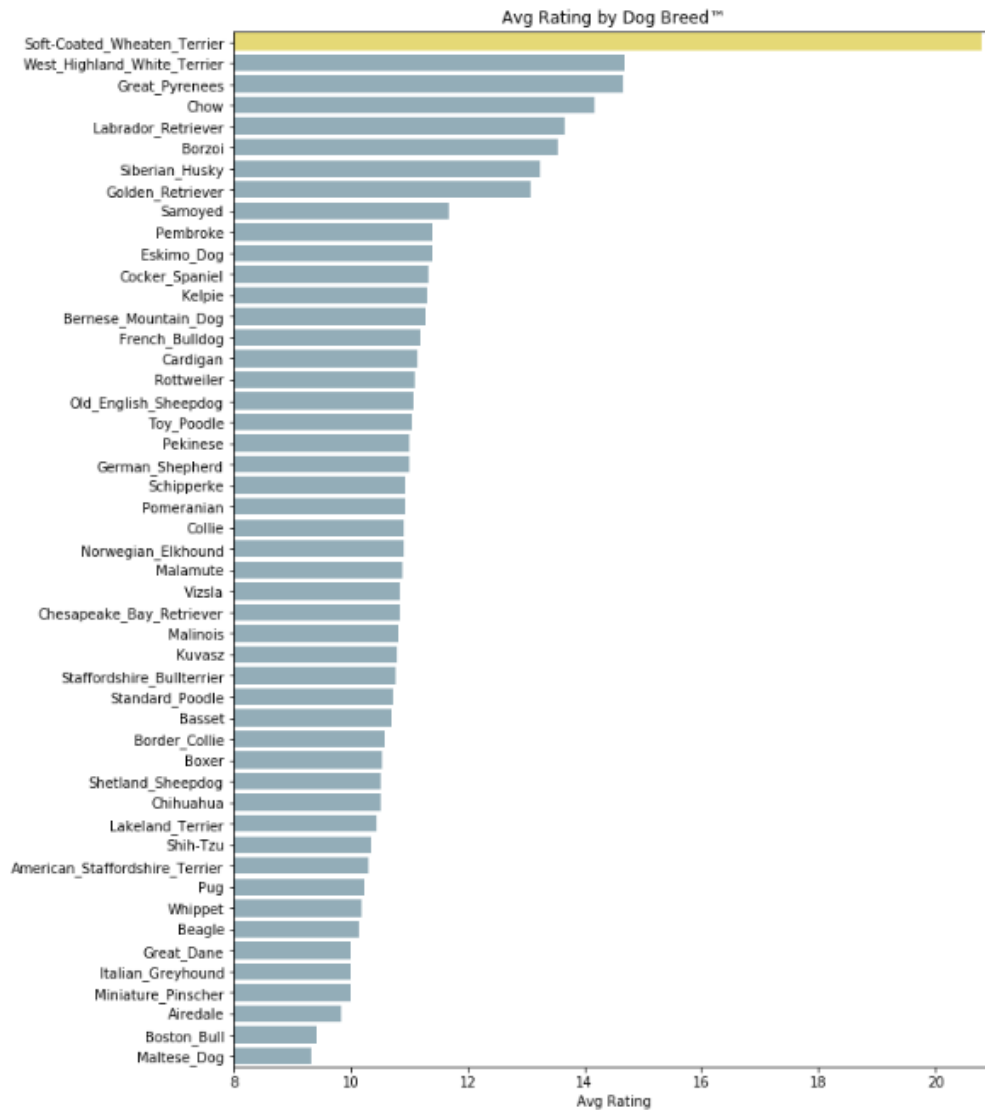
## Most Popular Dog Breed

We can see that the Golden Retriever dog breed received the most retweets and favorites from the tweets we have gathered. Almost 1.2 million favorites in total.



## Average ratings by Dog Breed

The soft coated wheaten terrier breed appears to have the highest average rating across all dog breeds in the data. 6 ratings more than the second highest: West highland white terrier.



### Most popular Dog(s) according to retweets & Favorites

The most popular dog was a Chihuahua named Stephan, according to the data collected. He received 121793 favorites and 57271 retweets on his tweet.

