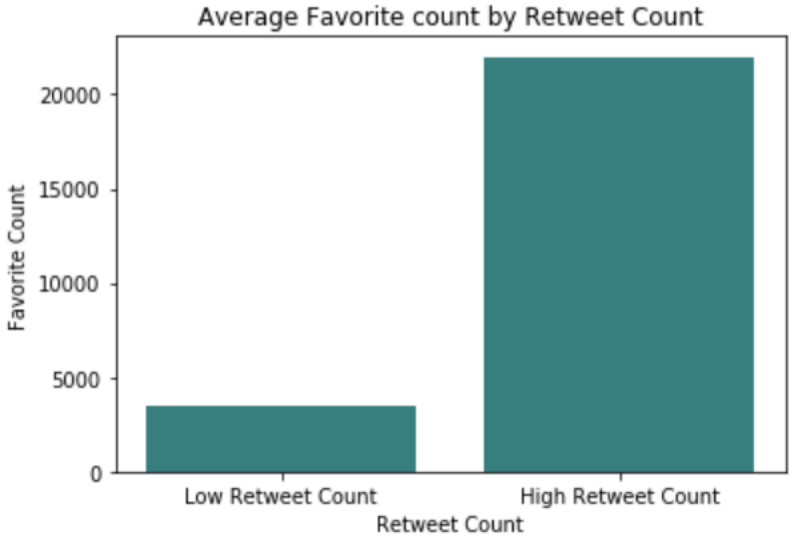
We Rate Dogs Twitter Analysis Report

I had the opportunity to dive deep and analyze some very interesting date coming from the WeRateDogs twitter profile. WeRateDogs post pictures of all types of dogs and rates the pictures, many of the pictures have very high ratings as most dogs are deserving of greater than 10 out of 10 in my opinion. After gathering, assessing and cleaning this data that was provided by Udacity, it was now time to analyze the data and see what insights we can learn about the dataset. I began the process by first analyzing correlations and data statistically and then after gaining some insight I began to dive deeper with visualization to help paint a clearer picture of those insights. I had four main insights from this data:

Analyzing correlation between Retweets and Favorites

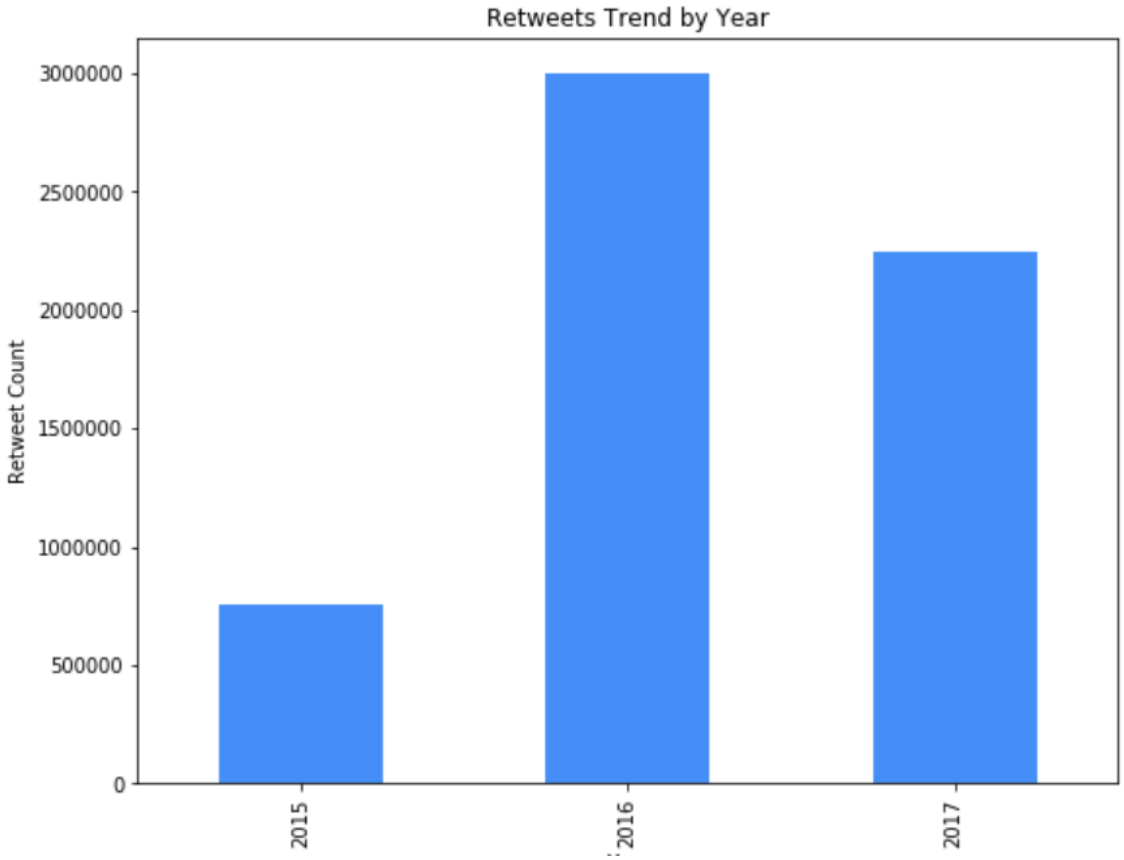
I wanted to see if there was a correlation between retweet count and favorite counts as I will be using these two variables to assess the rest of the insights. I did this by getting the average retweets and favorites. I then created a variable high and low for both favorites and retweets using greater than the mean and less than the mean. I Then created a bar chart plotting the average high and low retweet count on the x axis and the favorite count on the y axis. This chart showed that the high retweet count also had a high favorite count while the lower retweet count had a low favorite count. This showed me that there is a correlation between the retweet counts and favorite counts.



Analyzing Retweets and Favorites Overtime

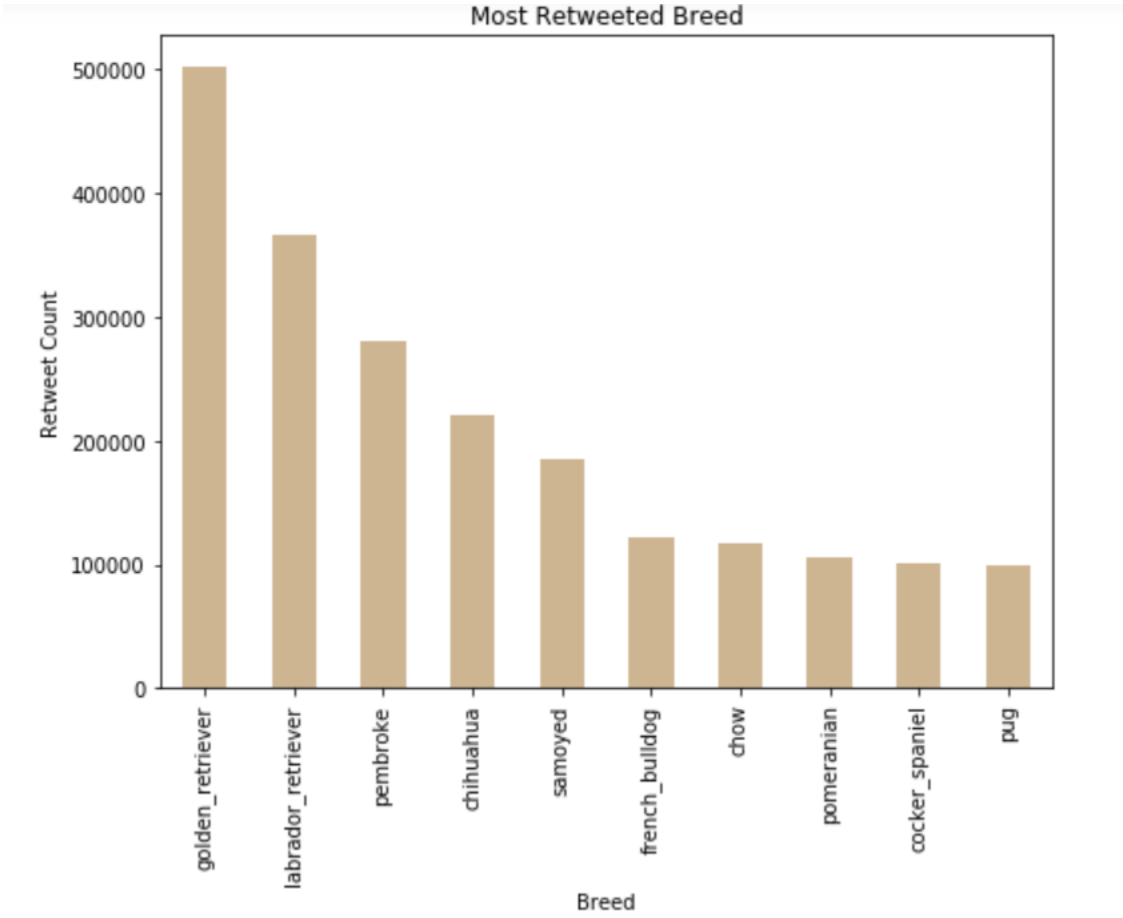
I wanted to analyze the amount of retweets over time to get an understanding of how there could be trends in retweeting. I did this by creating bar charts analyzing the retweet count by year and then again by month and doing the same for the favorites count. The posts dated from 2015 to 2017. I was able to get a very clear picture of how and when the retweets and favorites were happening the most. I learned that there were the largest amount of retweets in 2016 while the lowest amount was in 2015 and I also learned that the lowest amount of favorites also came in 2015 but the favorites gradually grew over time gaining more favorites on a year to year basis unlike the trend of retweets. I also analyzed these same variables on a monthly basis and got some interesting insight. I learned that both the favorite count and retweet count was very high in both January and

February, which I found to be interesting. When looking at the favorite count by month though it was clear that the top two months with the most favorite counts happened to be June and July with January and December coming in third and fourth. I found this to be interesting as the highest were in both the middle of summer and winter.



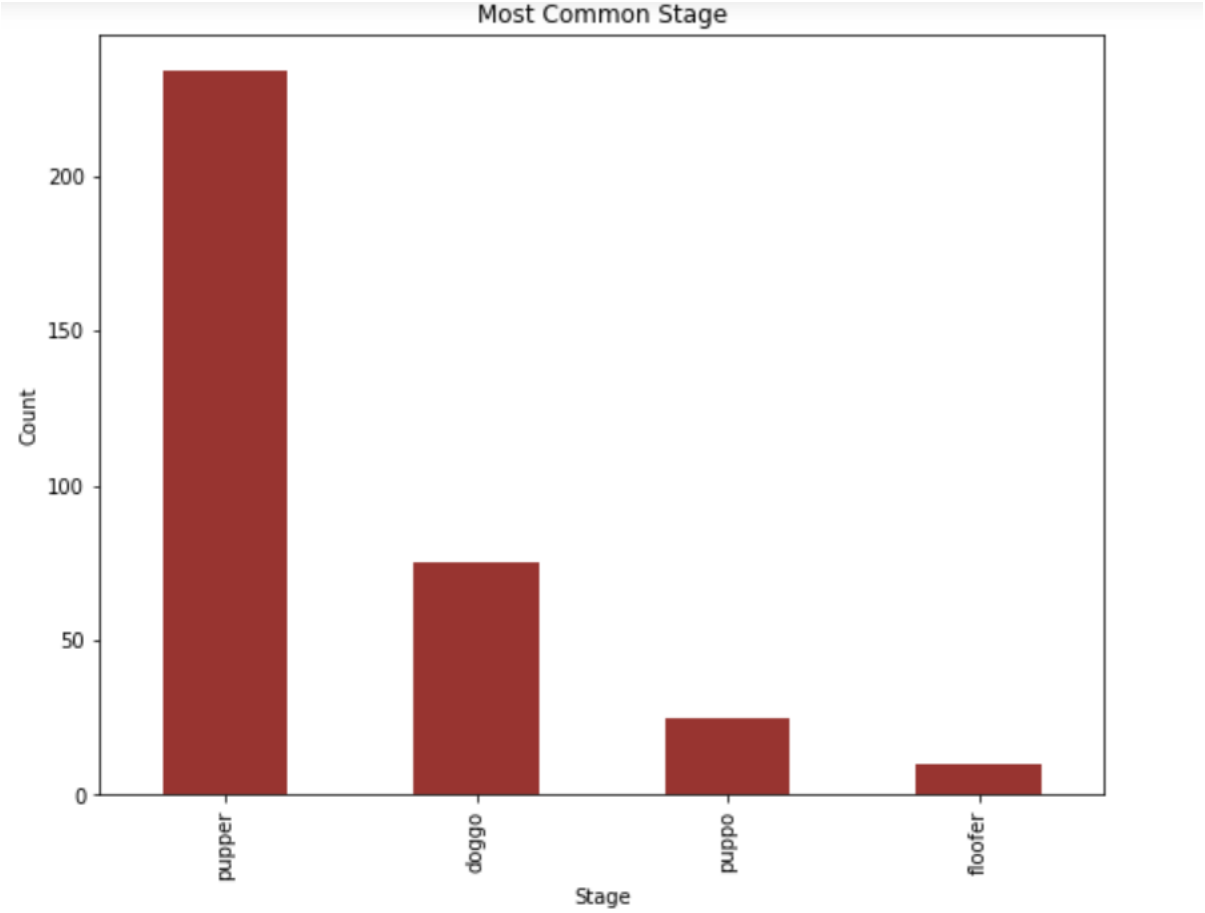
Analyzing most Popular Dog Breeds

Since all of this twitter data was about dogs I wanted to dive in and see which dog breeds were the most popular. I did this by analyzing statistically first and then creating visualizations for clarity. I plotted a bar chart showing the top 10 most retweeted dog breeds and then top 10 most favorited dog breeds. In both visualizations it was clear that golden retrievers were the heavy favorite landing at number 1 both times. Labrador retrievers also held strong at second in both favorites and retweets while the third-place breed was the Pembroke in both the retweeted and favorited, rounding out the three most popular dog breeds.



Analyzing the most common Stage

In WeRateDogs there are four stages tat a dog can be in: pupper, doggo, puppo and fluffer. I analyzed this statistically and visually, getting the value counts of each stage and then plotting this on a bar chart for visual clarity. The insight I got from this was that the majority of dogs were in the pupper stage while the least common stage was the floofer stage.



Conclusion

There is much more that can be learned from this data by analyzing this dataframe further but I was able to get some interesting insights in the analyzing and visualizing I was able to perform.