

# Report: Analyzing and visualizing data for the WeRateDogs tweet archive

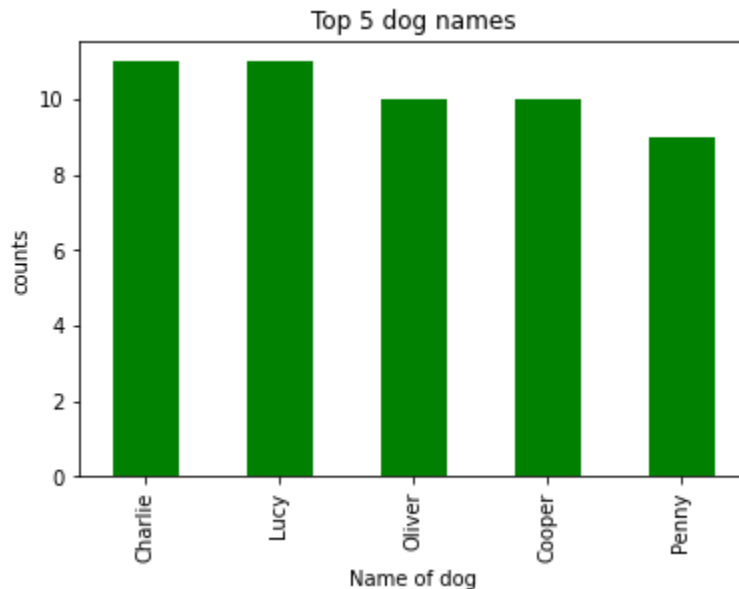
The analysis was carried out on the final data from the data wrangling step. The following analysis was done on the final master data

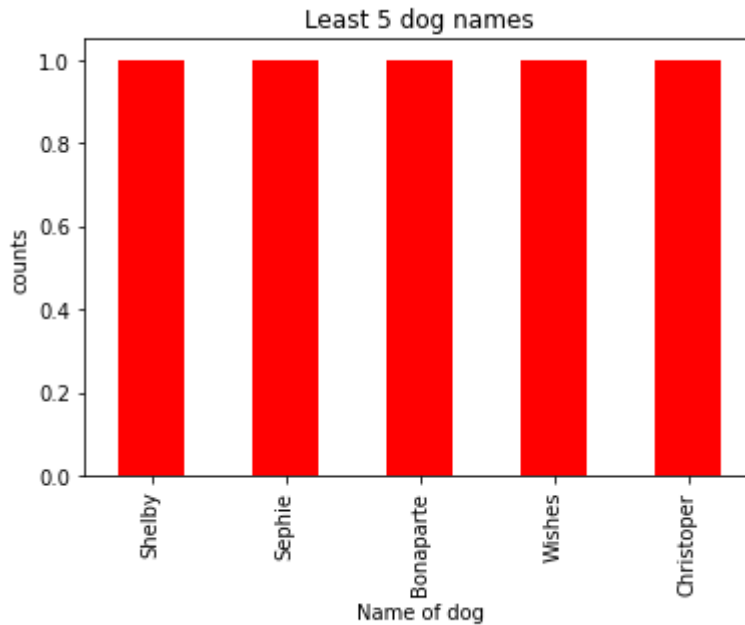
Insights:

- 1.What is the most and the least popular dog name?
- 2.What is the most commonly used Twitter Source
- 3.What is the rate of Weratedogs retweets over time?
- 4.What is the correlation between retweet and favorite counts

## 1.What is the most and the least popular dog name?

For the first section we extracted the top 5 and least 5 dog name from the data and used it to plot two different bar charts so as to properly visualize the most used and least used dog name. The charts can be seen below.

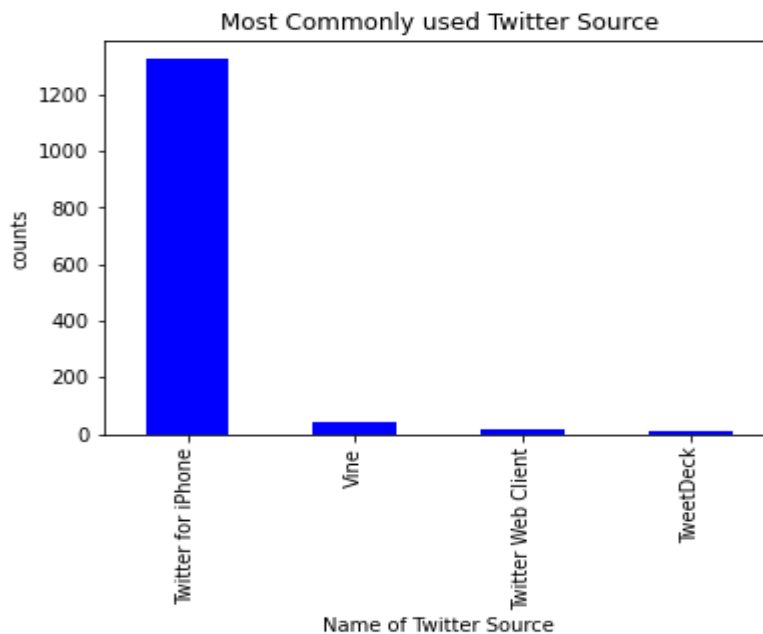




From the charts above we are able to see the most used dog names of which Charlie and Lucy tops the list and we have the least used names of which all 5 have the same occurrence. It would help dog owners choose names if they wish to base it on popularity.

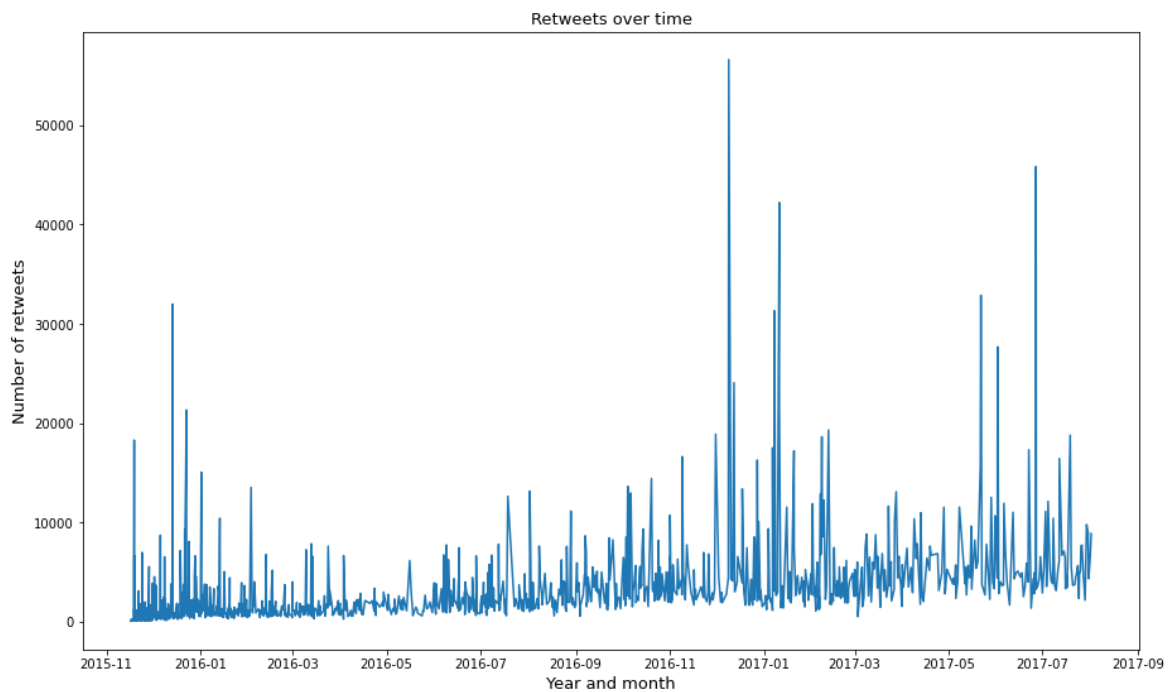
## 2. What is the most commonly used Twitter Source?

This insight here tells us about the most used Twitter source in the weratedogs dataset and to a large extent it helps us to know the platform most users tweet from.



### 3. What is the rate of Weratedogs retweets over time?

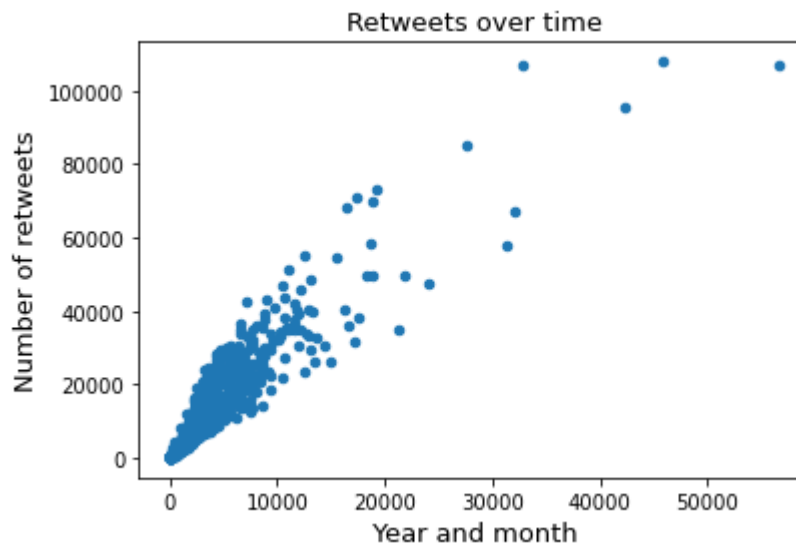
This is to show the number of retweets the Weratedogs tweets gets over a period of time. This is quite important because it helps us to gauge the popularity of Weratedogs tweet and also to know periods when there was a decline.



From the charts displayed we can see vividly that the month of December 2016 had the highest ever retweets on the Weratedogs tweets and between the periods of April and May 2016 we had the lowest number of retweets. We can see that the popularity of the Weratedogs is not a straight line growth.

#### 4. What is the correlation between retweet and favorite counts

The association between the retweet and favorite counts is the subject of the final realization. The scatter diagram between the two variables used for this is depicted below.



We can observe from this scatter plot that the number of favorites and retweets are positively correlated. Thus, it follows that the favorite count rises as the number of retweets does.

#### Summary of the Analysis

1. Charlie and Lucy tops the list of the most popular dog names with 11 occurrences each while we had about 5 names which are the least popular and this to a great deal implies that some names have a far reaching influence on choice of dog names than others.
2. The most commonly used Twitter source is Twitter for iPhone which implies that majority of the ratings come from users with iPhone.
3. The period that had the highest number of retweets on the Weratedogs ratings was in December 2016 while the period that had the lowest number of retweets was in May 2016. This can be a good way to gauge the popularity of the channel because every retweet increases the reach of the ratings tweet.
4. We can also see from the final chart that there is a positive correlation between the favorite count and the retweet count meaning the more the retweet the more likely it is to have a favorite for that tweet.