



**Act Report**

**A WERATEDOGS TWITTER DATA**

**Submitted by**

**Afyaa Alkhamisi**



## i. Overview

Twitter user @dog\_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

We got WeRateDogs Twitter data from their twitter archive from November 2015 to August 2017. We collected three sources of data for this project which are, the first is Enhanced Twitter Archive: contains very basic tweet information. The second source is Additional Data via the Twitter API: we gathered each tweet's retweet count and favorite ("like") count. The last source is Image Predictions File: is present every image in the WeRateDogs Twitter archive through a neural network.

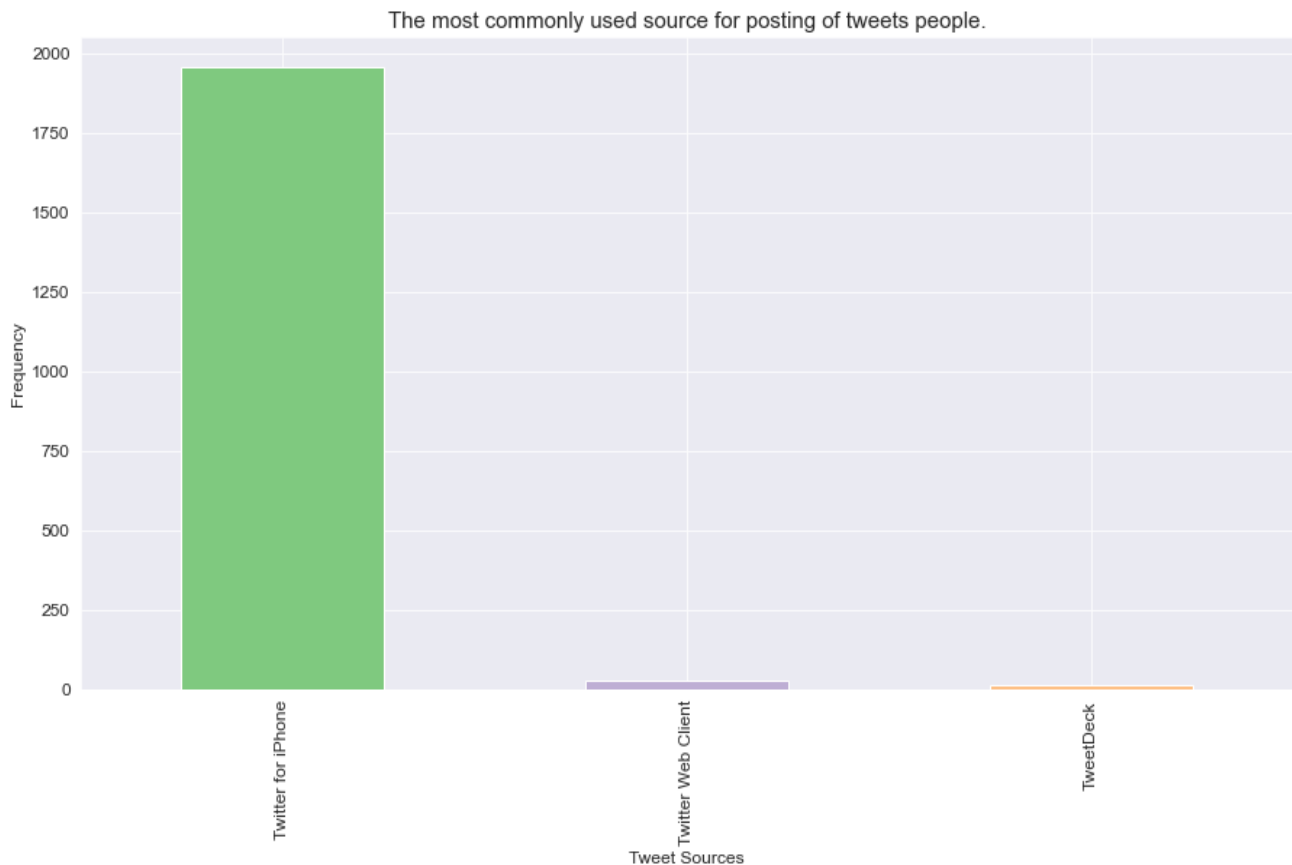
Therefore, to create interesting and trustworthy analyses and visualizations, we've decided to ask the following questions: What is the most commonly used source for posting of tweets people? Which are the most famous dog names? How have the effects of retweets and favorites count changed over time? Which are the most popular dog stages that people prefer? Which stage of those popular dogs has the highest average rating? What are the 10 most frequent predicted dog breeds?

But before we answer these questions, we wrangled these three datasets through assessing and cleaning them. The final dataset contains 1994 entries and 14 columns. Some columns have missing values, which we should keep in mind while looking at the visualizations below. Well, let's get started answering the questions!

## ii. Analysis and Visualizations

We have thought to ask a question to know the platform that people use to access to Twitter.

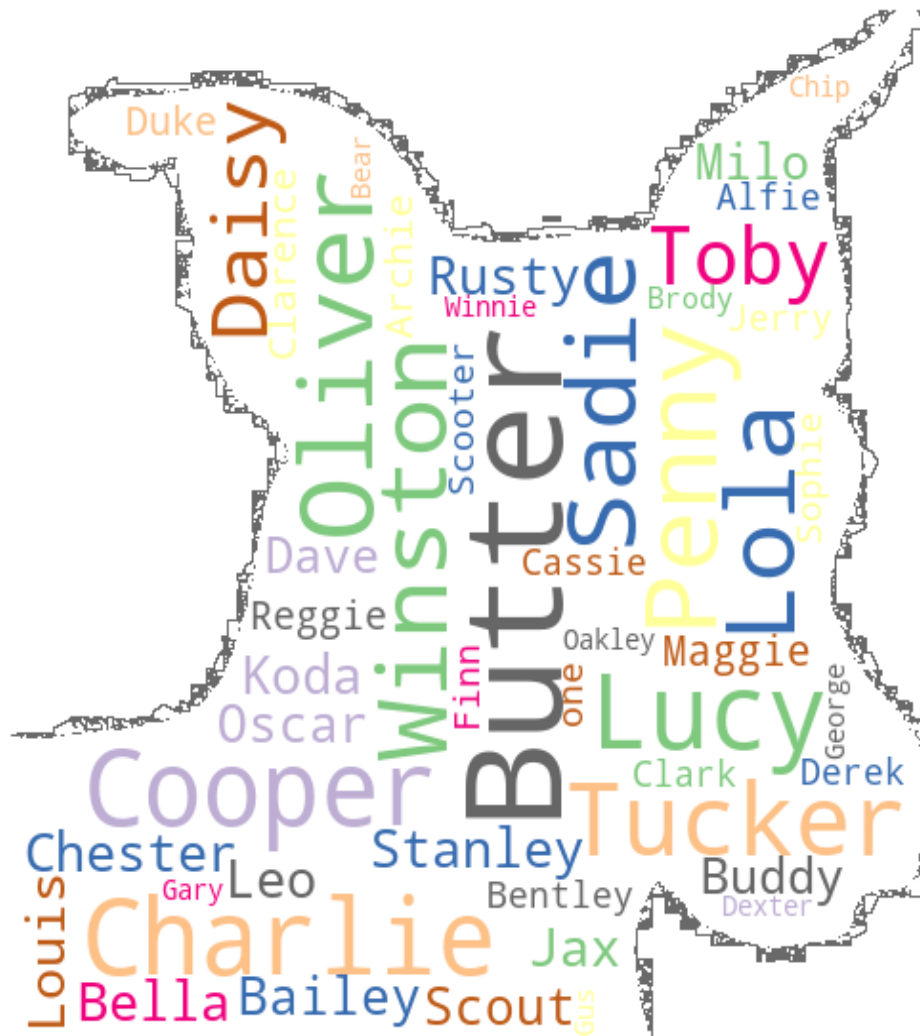
- What is the most commonly used source for posting of tweets people?



As shown in the above figure, the most commonly three used sources for posting of tweets people. First, it's "Twitter for iPhone" with 1955 counts, followed by "Twitter Web Client" with 28 counts, and "TweetDeck" with 11 counts.

## ii. Analysis and Visualizations

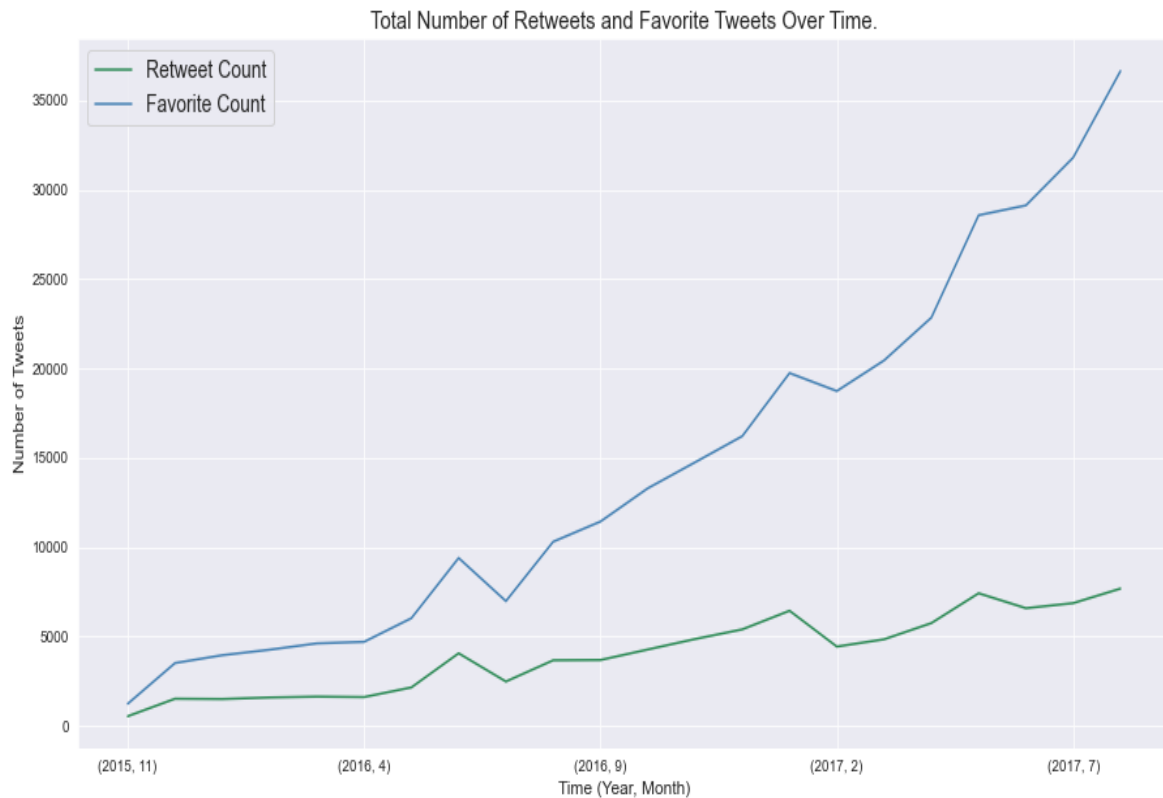
- Which are the most famous dog names?



To answer the question, we've built a word cloud showing the 50 most popular dog names. From the figure above, the most famous dog names were Charlie with 11 appearances of the name, followed by Oliver with 10, Lucy with 10, Cooper with 10, Tucker with 9, Penny with 9, Winston with 8, Sadie with 8, Toby with 7, and Lola with 7.

## ii. Analysis and Visualizations

- How have the effects of retweets and favorites count changed over time?

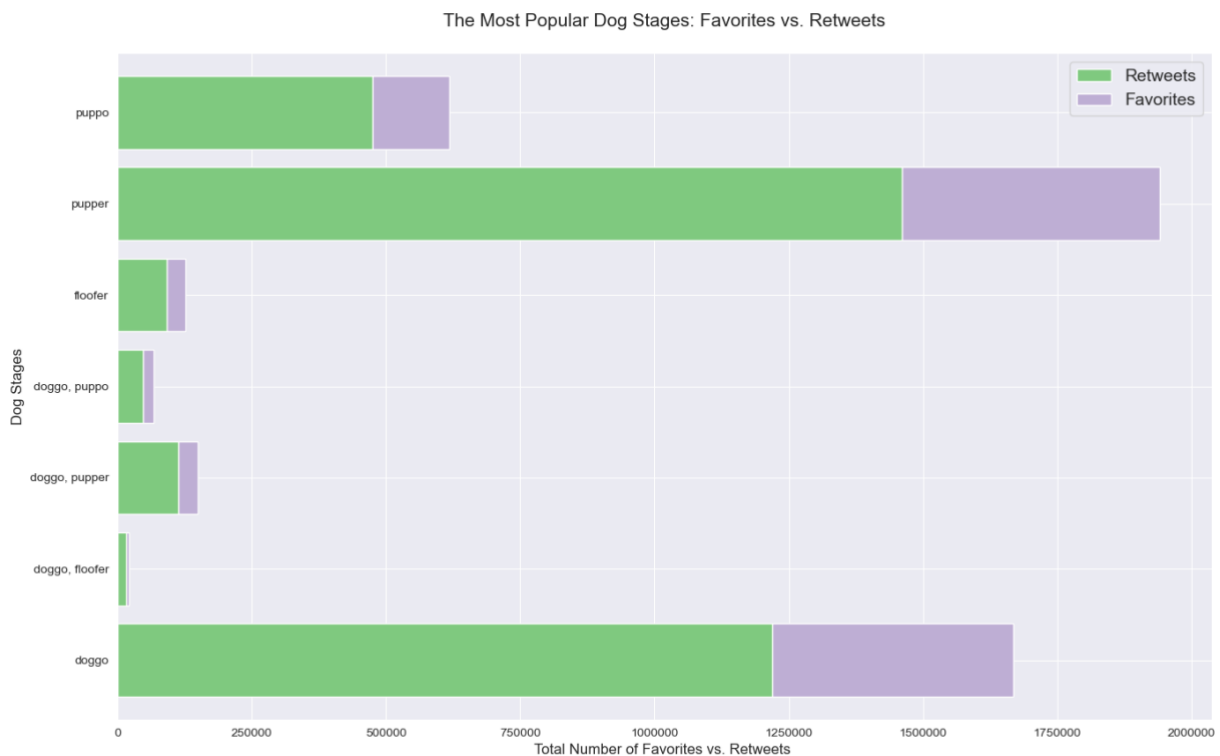


As shown in the above figure, the number of retweets and favorites has increased over time. Also, the number of likes was significantly more than the number of retweets. As it seems, in the year 2016 the number of likes and retweets decreased and then fluctuates over the months and years.

## ii. Analysis and Visualizations

Dogs have different stages: pupper, doggo, puppo, and floofer. We were wondering which of them occupied the most and least of dog stages in the dataset.

- Which are the most popular dog stages that people prefer?

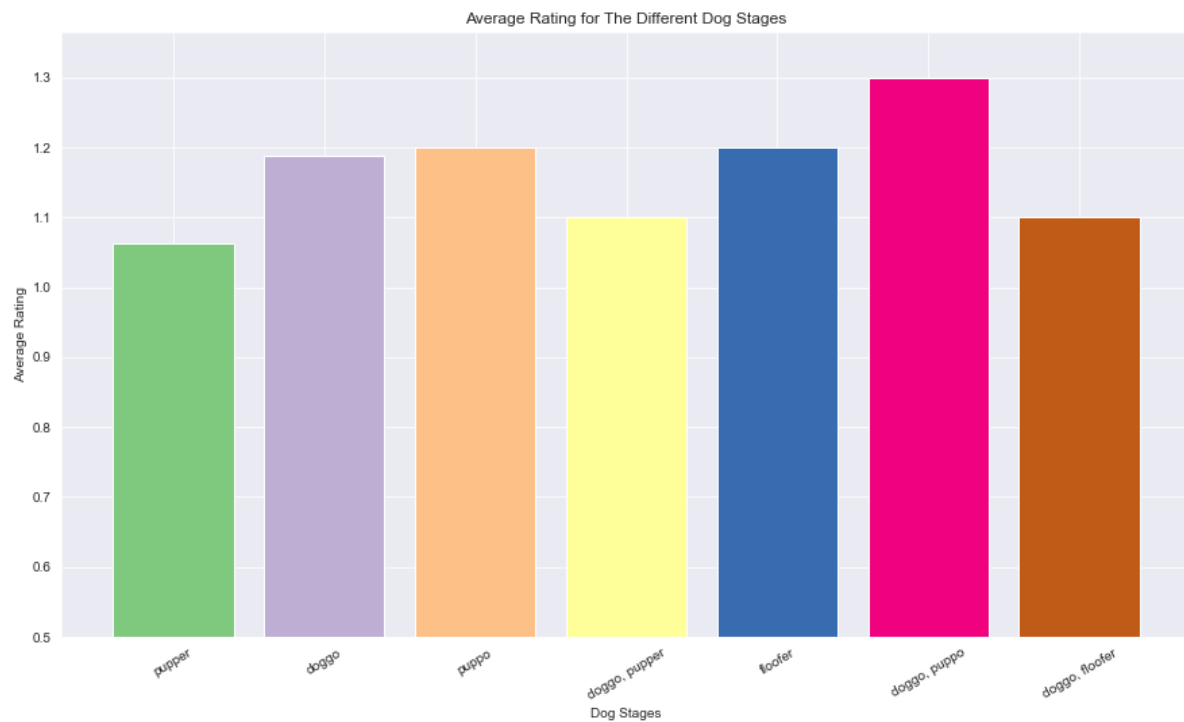


As we can see in above figure, the most common dog breeds preferred by people were "pupper" and "doggo", they got the most number of retweets and likes. And, The "doggo,floofer" breeds has the fewest number of retweets and likes.

## ii. Analysis and Visualizations

Among those the most popular dog stages, we were wondering who got the highest and lowest average rating

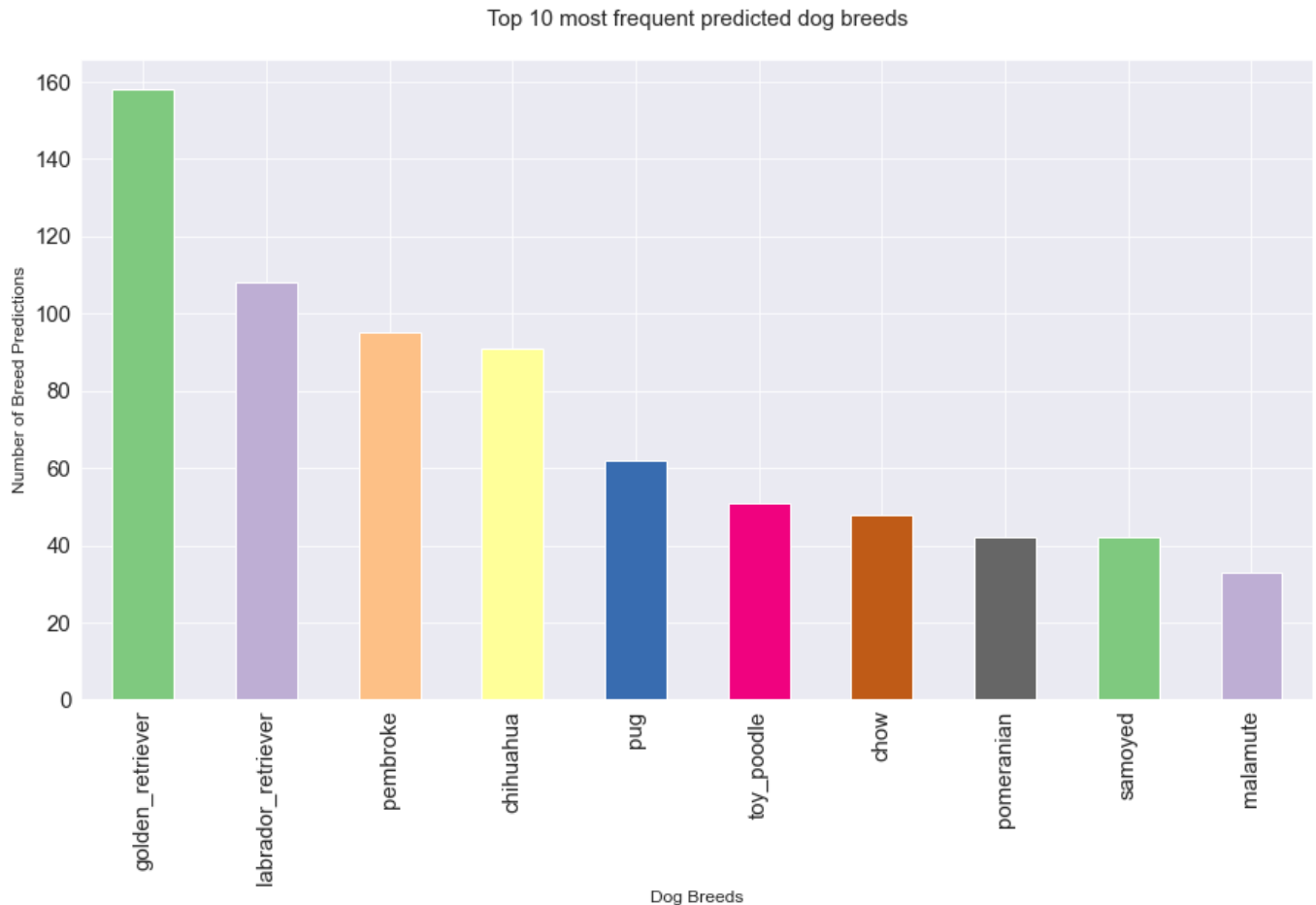
- Which stage of those popular dogs has the highest average rating?



From the above figure, the popular dog stage "doggo, puppo" has the highest average rating. Whereas dogs with the "pupper" stage the lowest average rating.

## ii. Analysis and Visualizations

- What are the 10 most frequent predicted dog breeds?



As we can see in above figure, the most frequent predicted dog breeds was golden\_retriever with 158 predictios, followed by labrador\_retriever with 108 predictios, and pembroke got 95 predictions. Then, chihuahua with 91 predictions, followed by pug with 62 predictions, toy\_poodle with 51 predictions, chow with 48 predictions, pomeranian with 42 predictions, samoyed with 42 predictions, and malamute with 33 predictions.

## iii. Conclusion

Briefly, we wrangled WeRateDogs Twitter data from their twitter archive and created interesting and trustworthy analyses and visualizations. In general, here we asked several questions and were able to answer them.