

# Act report

WeRateDogs is account in twitter that rate people's dogs, with rating value such as 14/10, 17/10.... etc. three files used *twitter\_archive\_enhanced.csv*, *image\_predictions.tsv* and *tweet\_json.txt*, was full of dirty and untidiness issues, so the beginning was with wrangling phase, this phase contain three main steps gathering, assessing and cleaning. All of this covered in *wrangle\_act.ipynb* file, after that all the three files merged into one data frame, *twitter\_archive\_master* data frame used for analysis and visualization.

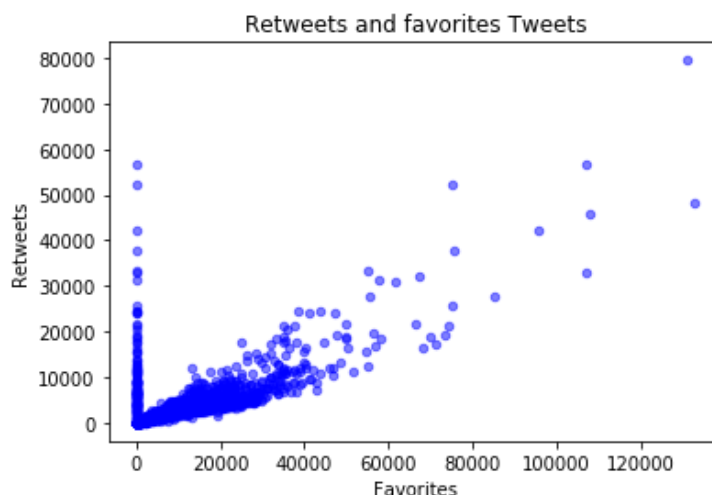
For the analysis phase, I selected four questions:-

1. Retweeted tweets are favorites and vice versa?
2. What is the most famous dogs' type?
3. What is the top dog breed?
4. What the most favorite dog breed?

## The insights

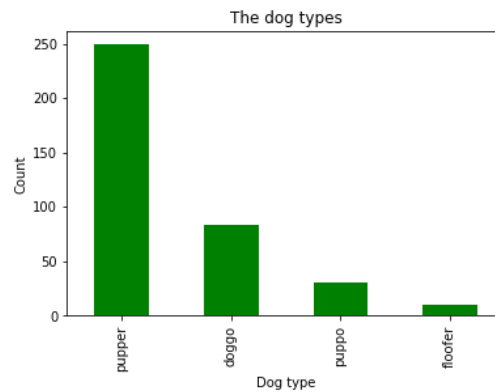
### Retweeted tweets are favorites and vice versa?

The relationship between the favorite tweets and the retweeted tweets, Plotting the favorite counts against retweet counts produced the diagram below. As shown in the figure, most dots are in the zero value, this reflects that the retweeted tweets are not favorite tweets and vice versa. This relation seems normal because who favorite a tweet not necessarily will retweet the same favorite' tweet.



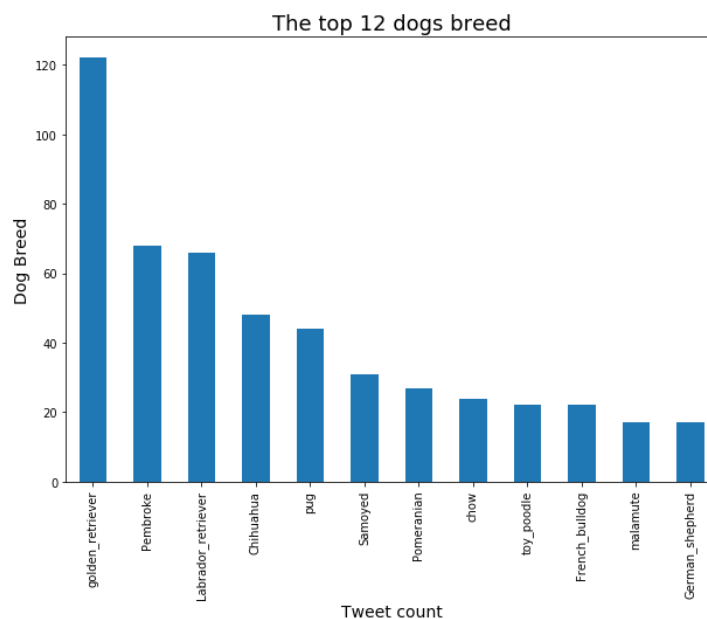
## What are the most famous dogs type?

In the data frame, there's four main dogs type, as shown in the figure the most in twitter\_archive\_master file is pupper type, that reflect most the dogs in the data frame goes under pupper type, while less than 50 goes under floofer type.



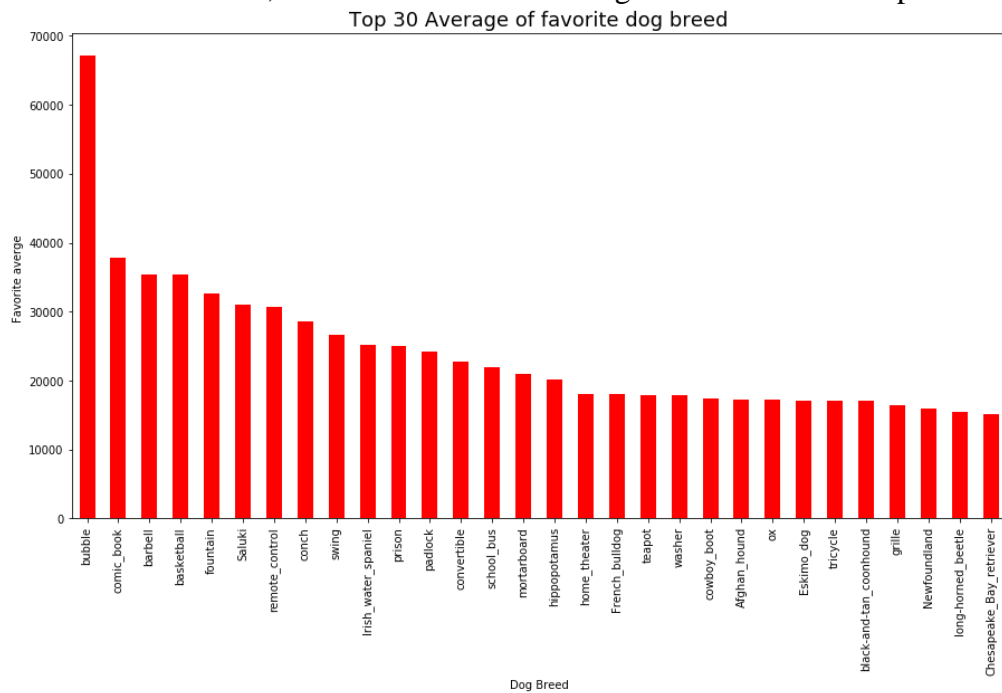
## What are the most famous dogs type?

So many dog breed mentioned in the data frame, this diagram show the top 12 famous dog breed in twitter\_archive\_master file, as shown in the figure, golden retriever varies than the rest, where it have more than 120 tweets, the rest are close to each other where they have less than 80 tweets.



## What the most favorite dog breed?

The top 30 of favorite tweets, the bubble is the highest bar, which means the bubble is the dog most favorite on twitter, others close to each no huge difference as the top one.



This dataset still full of issues even after the wrangling phase, due to the short time, I covered the top issues and as mentioned 8 at least.