

# DATA ANALYSIS AND VISUALISATION REPORT BY ISHAN ARORA

## Introduction:

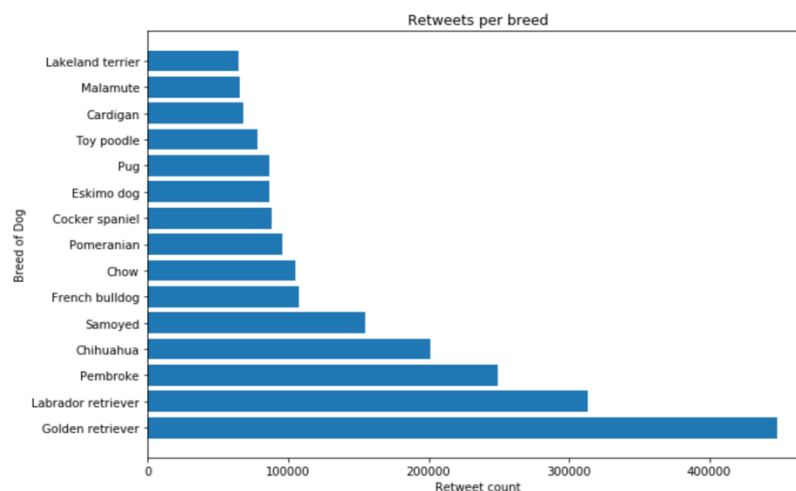
WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "[they're good dogs Brent](#)." WeRateDogs has over 4 million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

After thoroughly cleaning this dataset we will gather various insights about this data with help of some visualizations too

## Q1) Breed of Dog which receive more retweets?

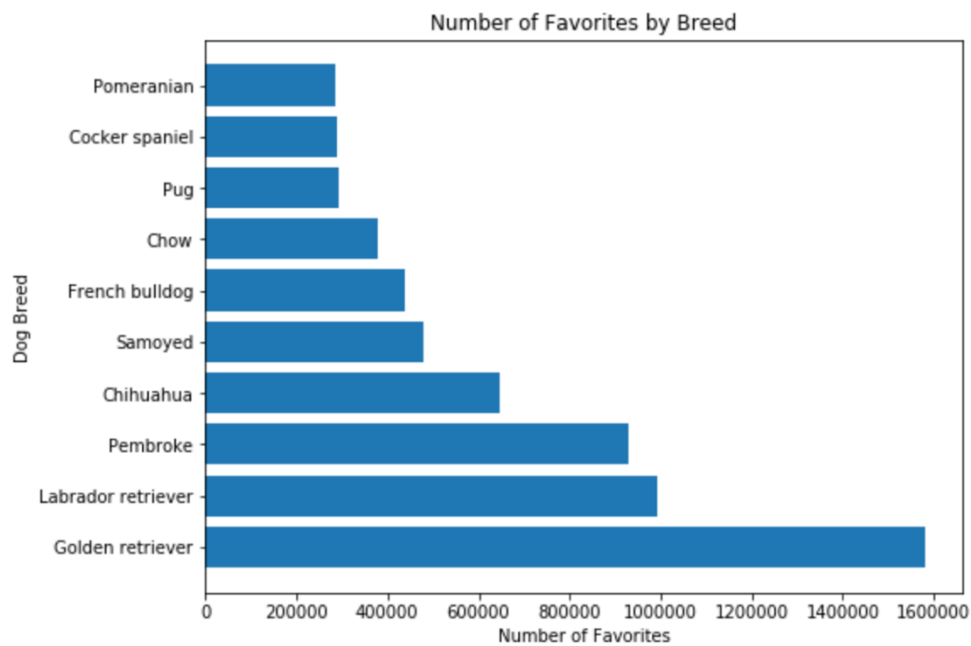
We found that Golden retriever was the one with maximum retweets.

Also most of the retweets were in the range 100k to 500k

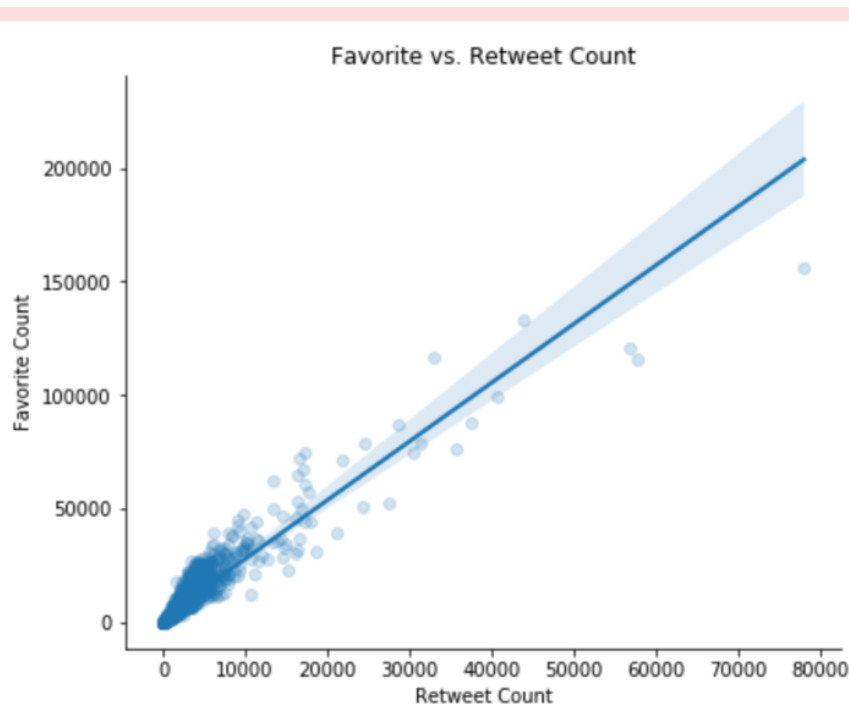


## Q2)What are number of favorites by the breed

Again in this list too Golden Retriever topped the charts and we found out that every breed has at least 200k favourites peaked by 1.6million



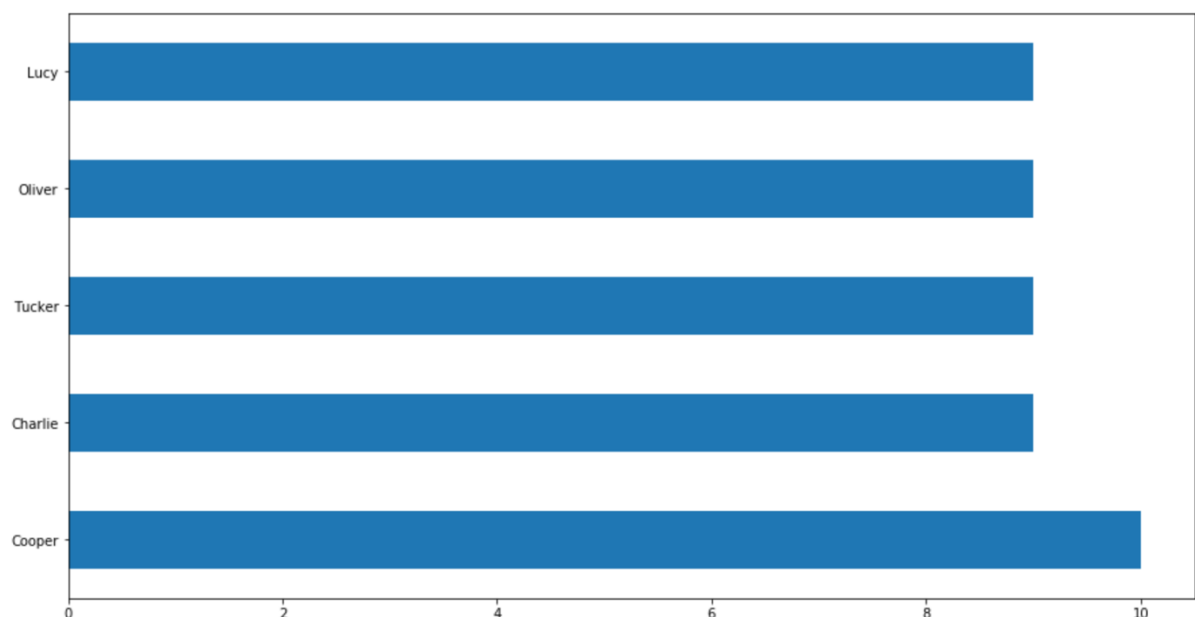
## Q3)Is there a correlation between favorites and retweets?



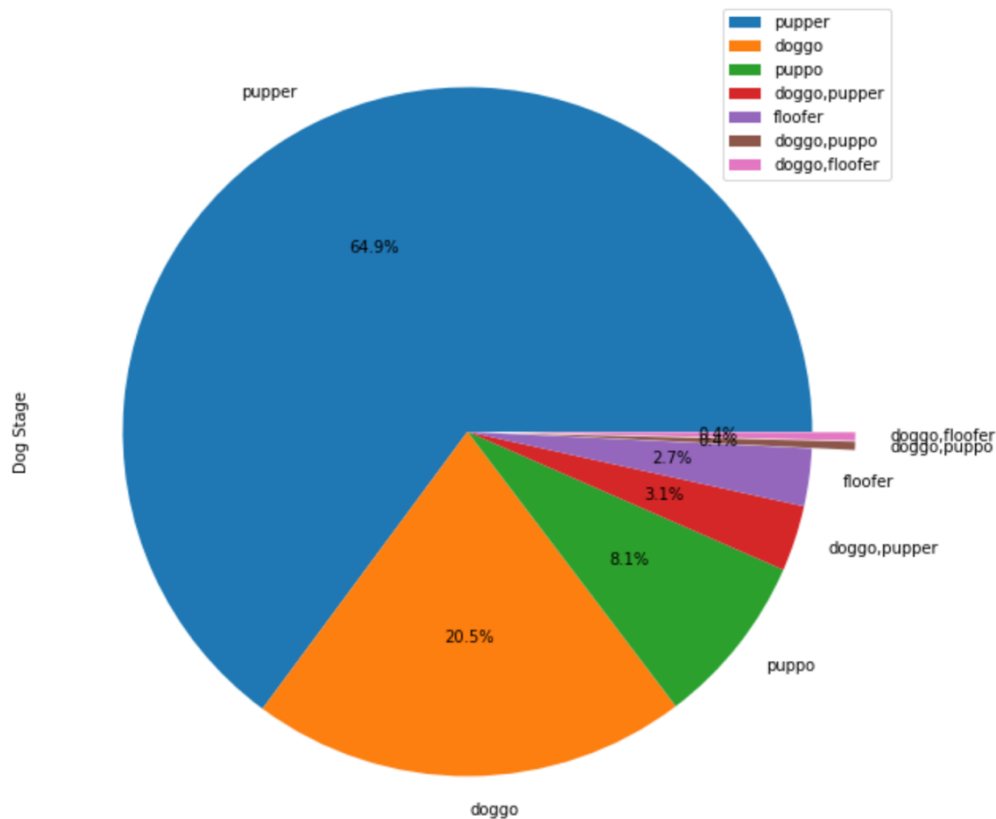
We find out that there is indeed a correlation between favorite breed and number of tweets retrieved.

#### **Q4) What are most popular dog names?**

We found out that Cooper is the most widely used dog name out of the name data that was available. Although the dog name data was a bit less so this might be a bit inconclusive



## Q5) What are Proportion of various dog stages?



I first removed the values and then created a pie chart to see what was the distribution of dog\_stages among the WeRateDogs' Dog population

- More than half of it was covered by puppies: These are the dogs which are still in training hence a bit immature
- 1/5th of the population was covered by the more mature Doggo population