

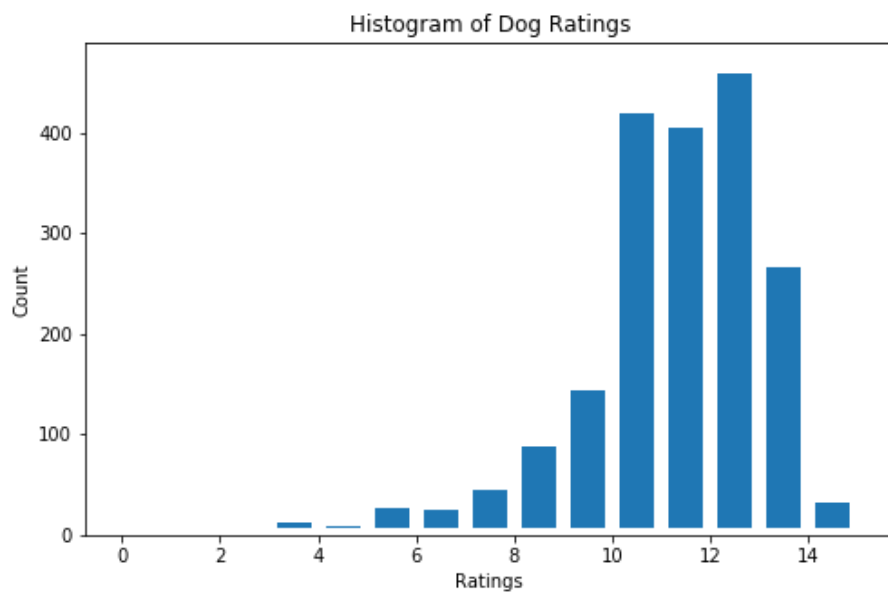
## ACT REPORT

In previous sections of this project, I used Python and its libraries to structured and clean dirty data into the desired format for better analysis and visualizations.

For the fourth and last section of the project, I undertook analysis and plotted visualizations about four data aspects :

### 1. A Rating Analysis

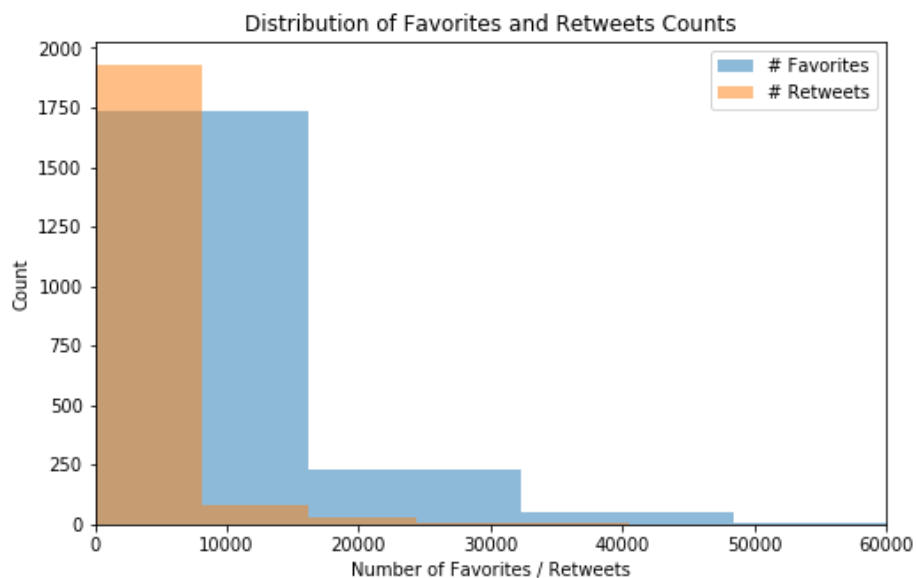
I plotted a histogram to figure out how the dog's ratings are distributed



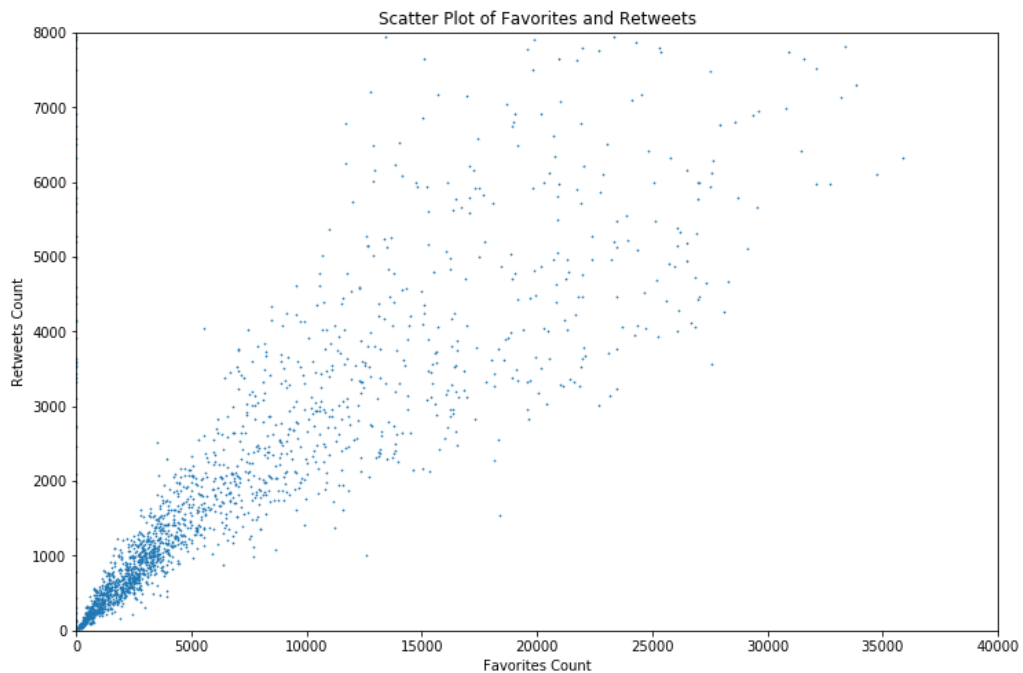
The plot shows a heavily skewed histogram to the right. It seems that the vast majority of dogs are rated between 10 and 15.

### 2. Favorites and Retweets Analysis

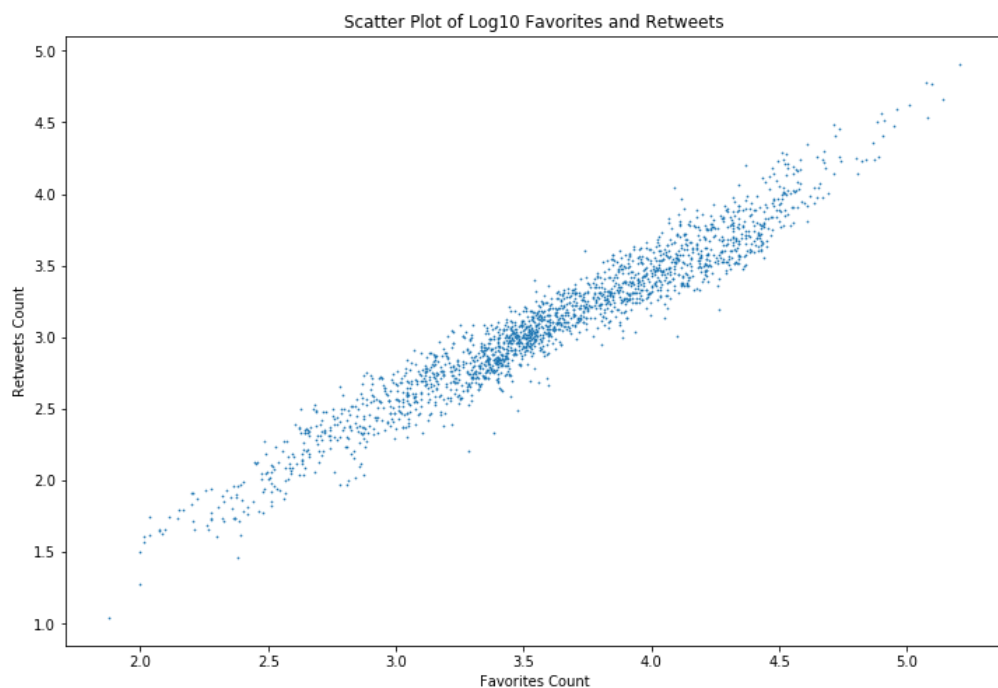
I plotted two histograms to figure out how the dog's retweets and favorites are distributed:



The two histograms show the same behavior : Distribution of favorite\_count and retweet\_count follow the same pattern. Let's check this with correlation

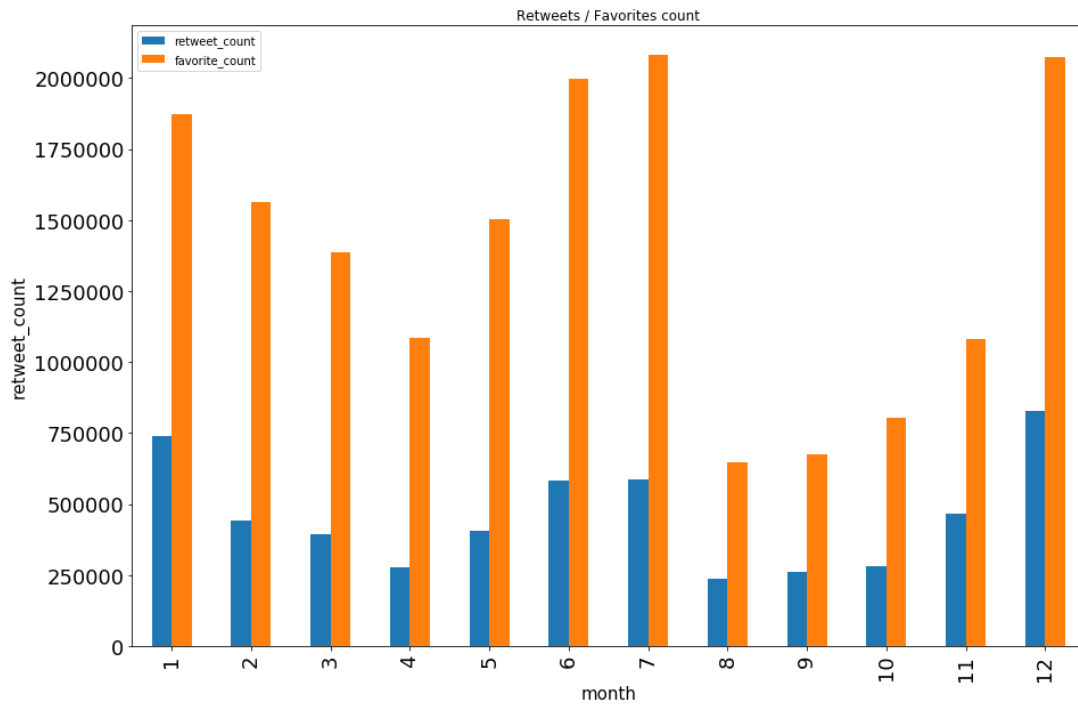


The scatter plot shows a pretty good correlation . Let's see if logarithm 10 will smooth this correlation.

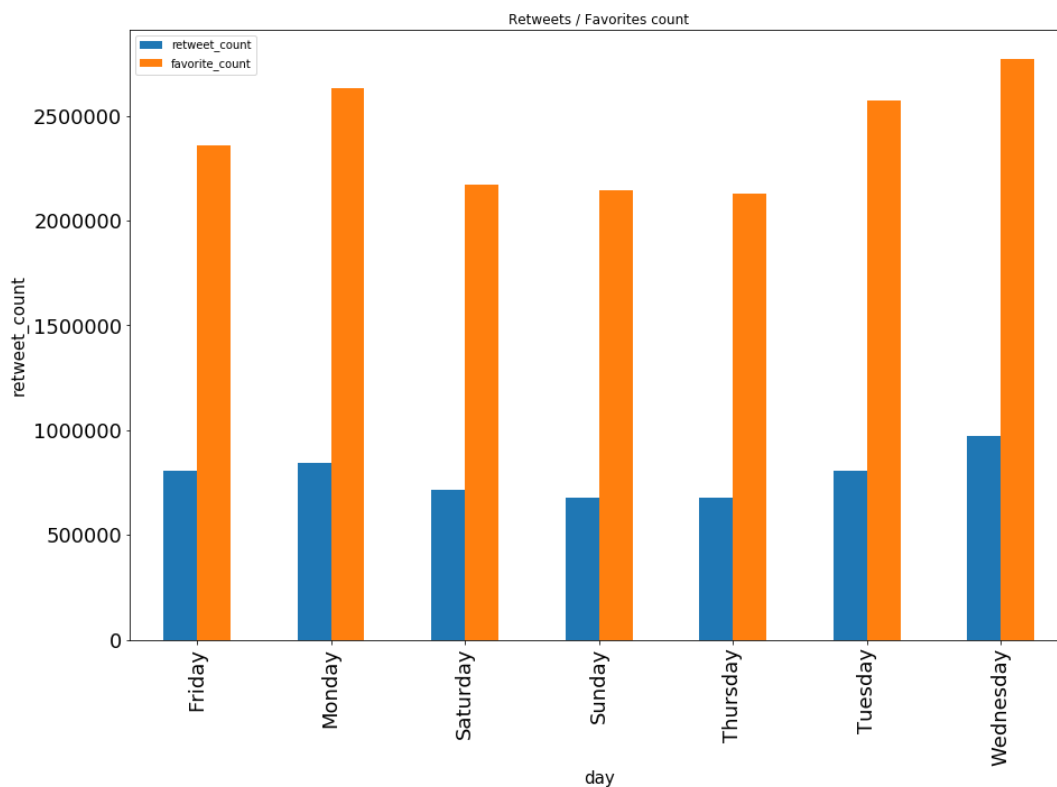


We can clearly conclude there is a strong correlation between favorites and retweets. That make sense since People who favorite seem to almost always retweet.

After this I analyzed retweets and favorites by month and days



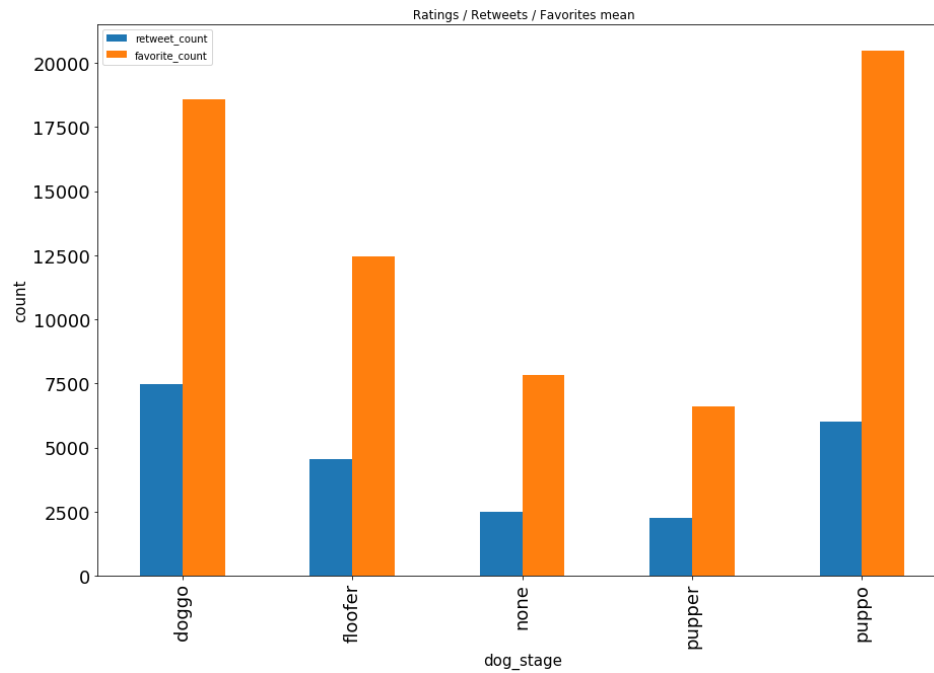
The plot bar chart shows :The highest number of retweets were in January and December and The highest number of favorites were in July and December.



The highest number of retweets and favorites both were on Wednesday.

### 3. Stage Analysis

In this part I asked the question: What is the dog stage which received the largest number of retweets and favorites in average ? to answer this question I plotted the following bar chart

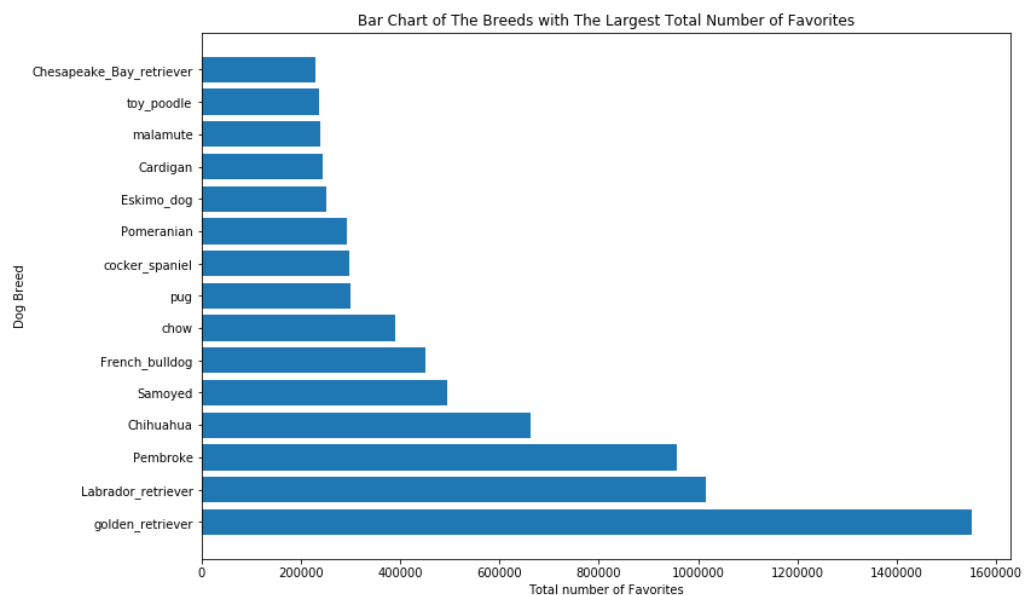


We can fairly say that 'puppo' and 'doggo' received the largest number of retweets and favorites in average

#### 4. Dog Breed Analysis

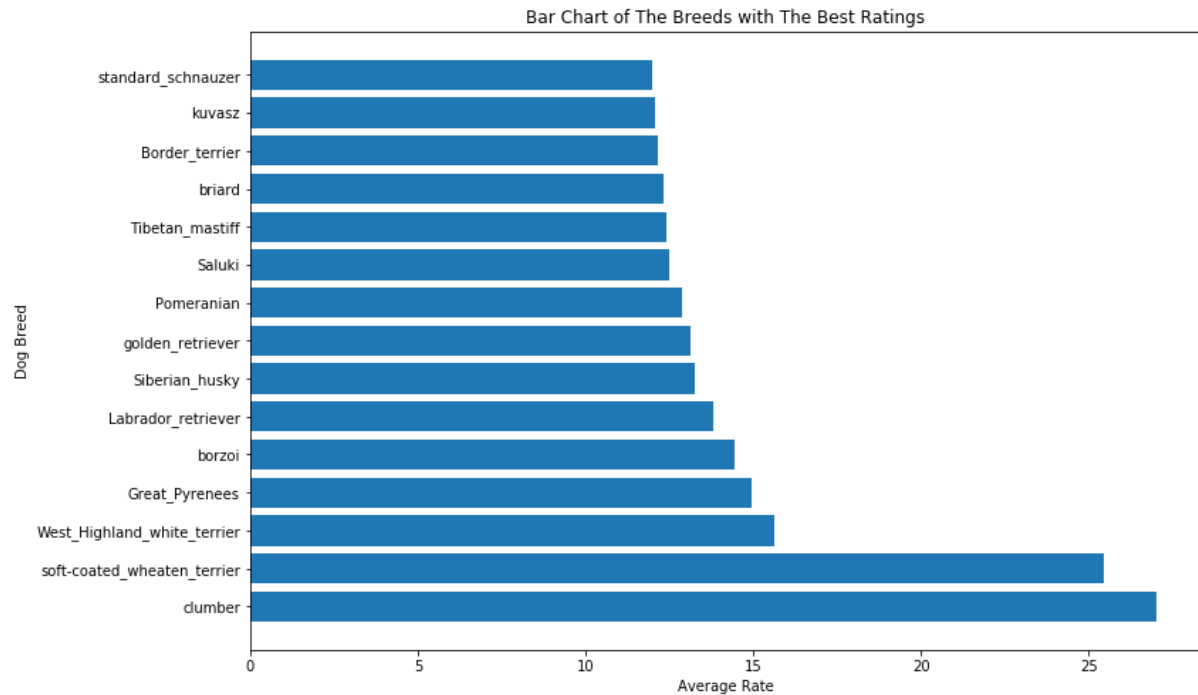
I merge the tweet and images datasets to be able to get insights about dog breeds vs ratings/favorites

This horizontal Bar chart shows the most 'favorites' breed.



The **Golden\_retriever** is the breeds that have the largest number of favorites, followed by the **Labrador\_retriever** and **Pembroke**

This horizontal Bar chart of best rated breeds.



The **Dumber** and the **Soft\_coated\_wheatern\_terrier** are the breeds with the best ratings (above 25 in avergae).