WeRateDogs Twitter Archive - Act Report

By-Piyush Sinha

WeRateDogs Data

The WeRateDogs Enhanced Twitter archive contains data extracted from 2356 of the 5000+ tweets from the @dog_rates twitter account, posted between November 15, 2015 and August 1, 217. This data is comprised of dog ratings that were taken from the text of the tweet along with the dog name and dog stage if present.

The retweet count and favourite count for each tweet were not included in the enhanced archive, and so I had to download this additional data from the twitter account using the tweet ID from the archive.

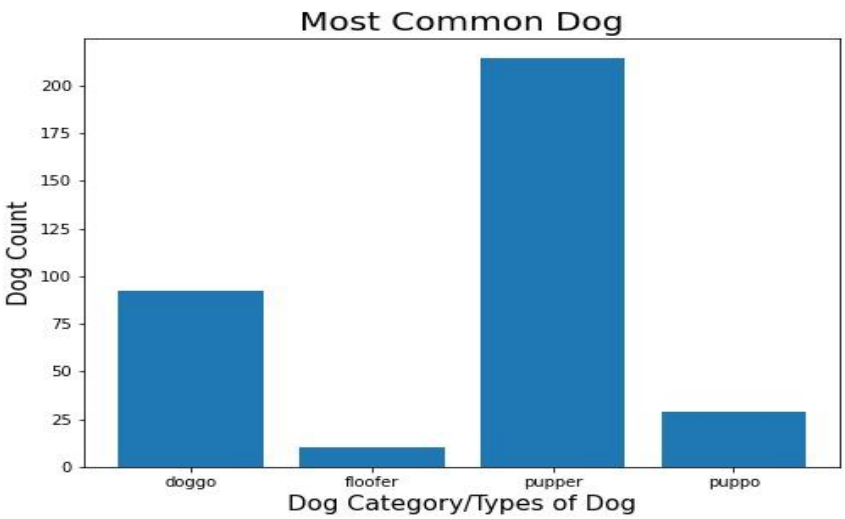
Along with the Twitter data, I also downloaded an image predictions file from Udacity servers containing the top 3 predictions for dog breeds based on the images from the tweets (there can be up to 4).

Insights

1.Most Common Dog

By performing visualizations on various dog category, we came to know that the most common or we can say most loved is: - **PUPPER**

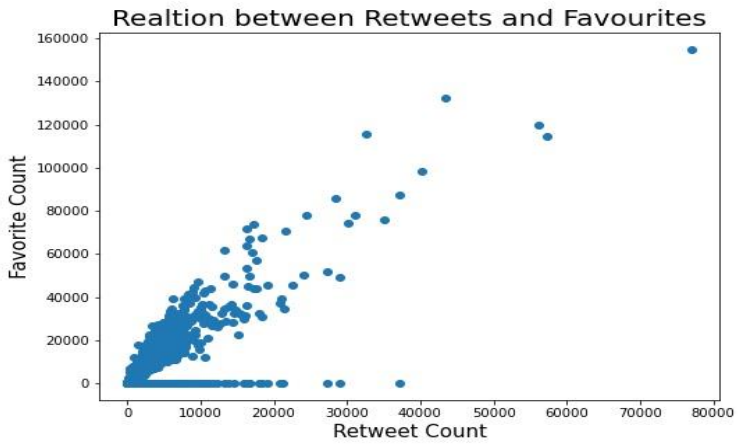
This was one of the most easy but valuable insights we got from this data.



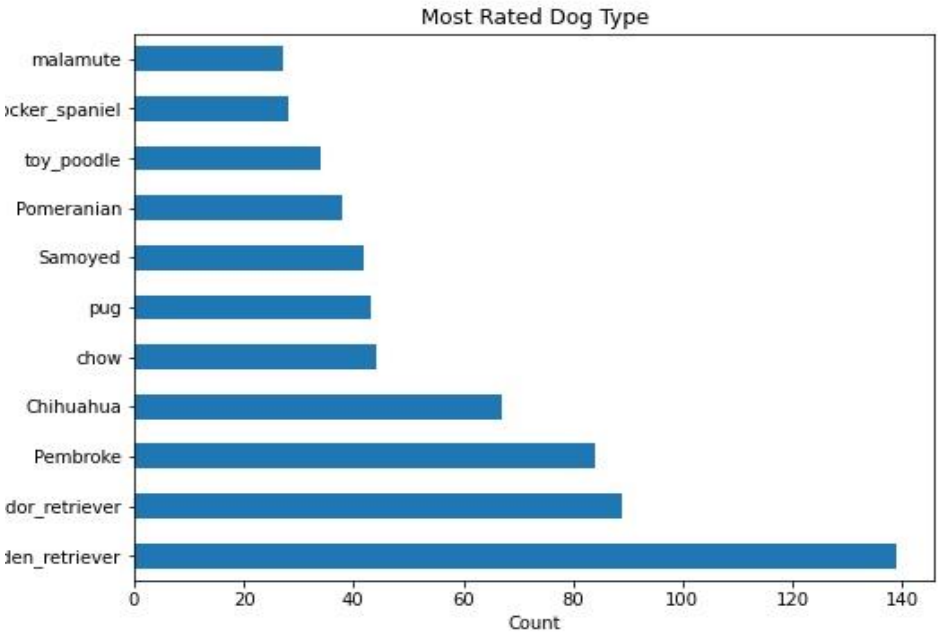
2. **Relation between favorites and retweets**

According to the insight generated in the jupyter notebook we can say there is a strong positive correlation between favourite_count and retweet_count which is 0.79677 **.**

This strong positive corelation can be analyzed with the help scatter plot also.



There is a strong relationship between a tweet's retweet counts and favourite counts; in other words if a tweet has a high favourite count it will also have a high retweet count. Plotting these two values against each other produced the plot below, and clearly shows a high correlation by the very tight collection of points along the angled, straight line.

3.Breed Prediction: MOST COMMON BREED

From the insight generated the most loved breed is **GOLDEN RETRIEVER**