Wrangle and Analyze Data @WeRateDogs

Project Analysis Report by Morayo Egbewumi

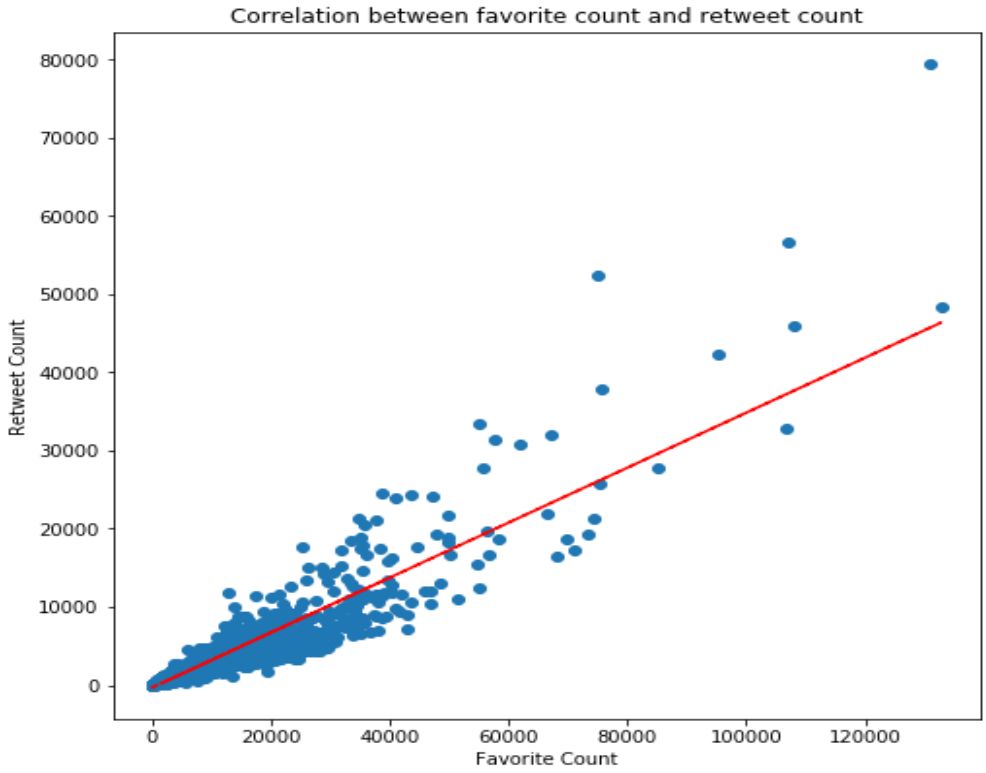
Introduction

WeRateDogs is a Twitter account where users may share images of their dogs and get ratings from the account's followers. Therefore, in this project, both the tweet data and the supporting data have been supplied for us to wrangle (clean and convert) and get ready for analysis. Once we have organised our data, we may utilise it to get some findings. In this paper, we show how to leverage our gathered data to provide some practical insights.

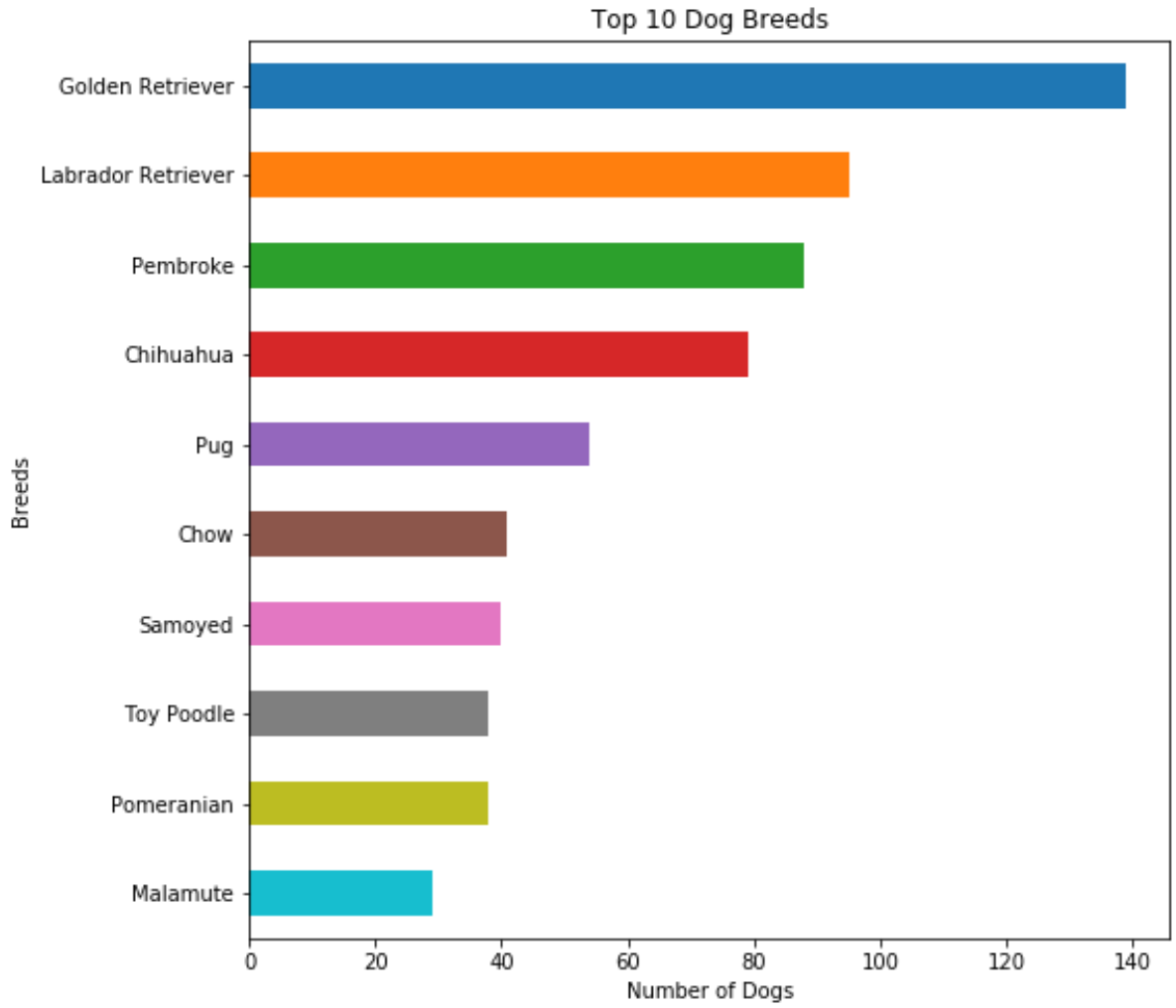
Analysis

Let's first examine the relationship between the quantity of favourites (likes) and retweets for each tweet. We can see that there is a positive connection between these two variables from the scatter plot below. That is, individuals frequently retweet tweets they find interesting. Further examination of our data reveals that tweet #348, which has 132,810 likes, is the most well-liked tweet. However, with only 52 likes, tweet #1116 has gotten the least amount of attention.

Additionally, according to our data, tweet #247 has received zero retweets, whereas tweet #1116 has received two total retweets.

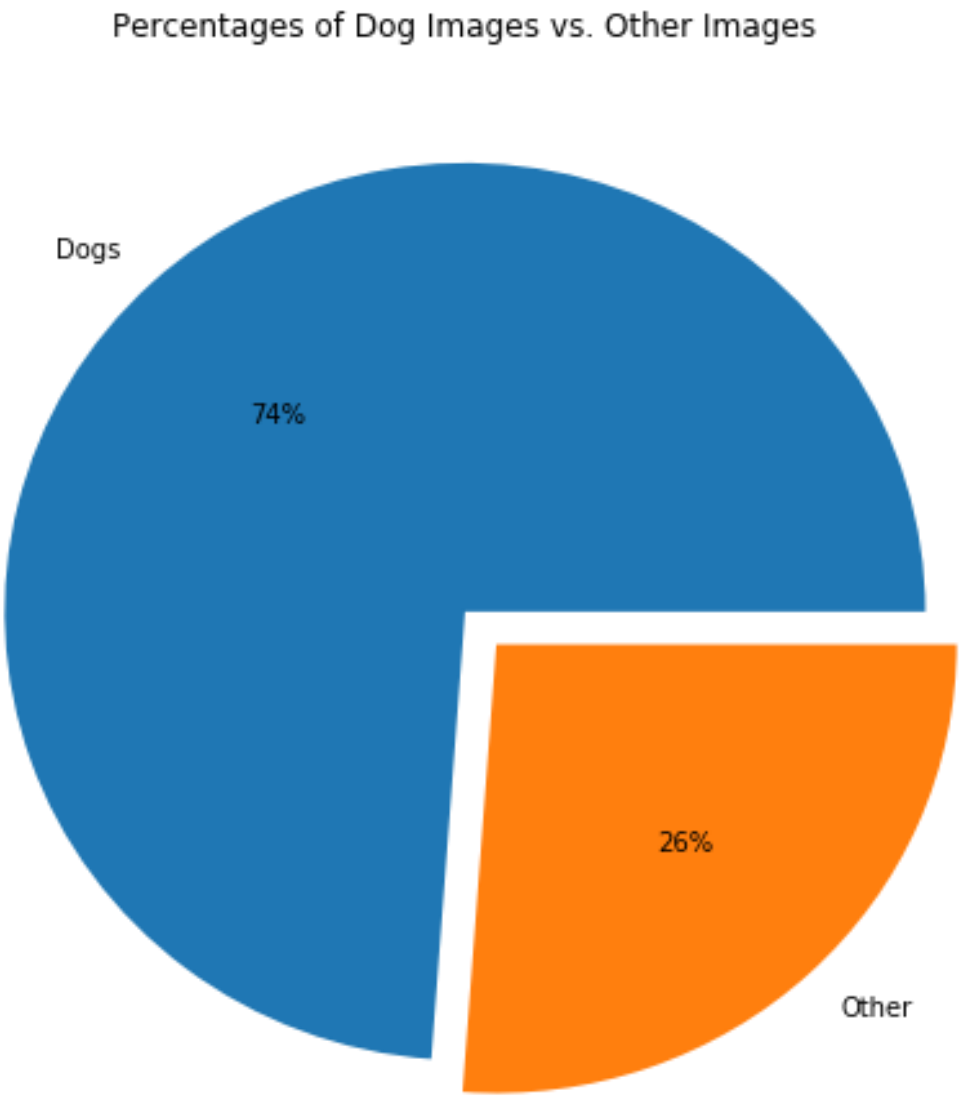


As can be seen in the horizontal bar graph below with 139 tweets, Golden Retrievers are the most popular dog breed in our sample, followed by Labrador Retrievers with 95 tweets.



The data also reveals that Puppers, which made up 224 tweets in our sample, are the most popular dog stage.

According to the statistics, Charlie and Lucy are the most popular dog names, with each name appearing in 11 tweets (Lucy for female dogs and Charlie for male dogs).



Finally, combining the archive table with the prediction table gives us a better understanding of the accuracy of our prediction system. The pie chart above, for instance, indicates that just 74% of the photos in our dataset are of dogs. However, it became clear that the algorithm misrepresented several dog photographs as other things when we looked at some of the 26% of data that were left. As an illustration, the classifier treated a dog in a shopping cart like a shopping cart rather than a dog.

With this knowledge, we can tell that the current image classifying method may benefit from some enhancement.