

Wrangle and Analyze Data

Act report

INTRODUCTION

The dataset that we will be wrangling and analyzing and visualizing is the tweet archive of Twitter user @dog_rates, also known as WeRateDogs. In this report I took data from the twitter account "WeRateDogs" which tweets about dogs types and rate them, the rate are from 10 points but most of times they give thing more than this like 12/10, first I did gather the data then assess the data then I cleaned the data, then I analyze the data.

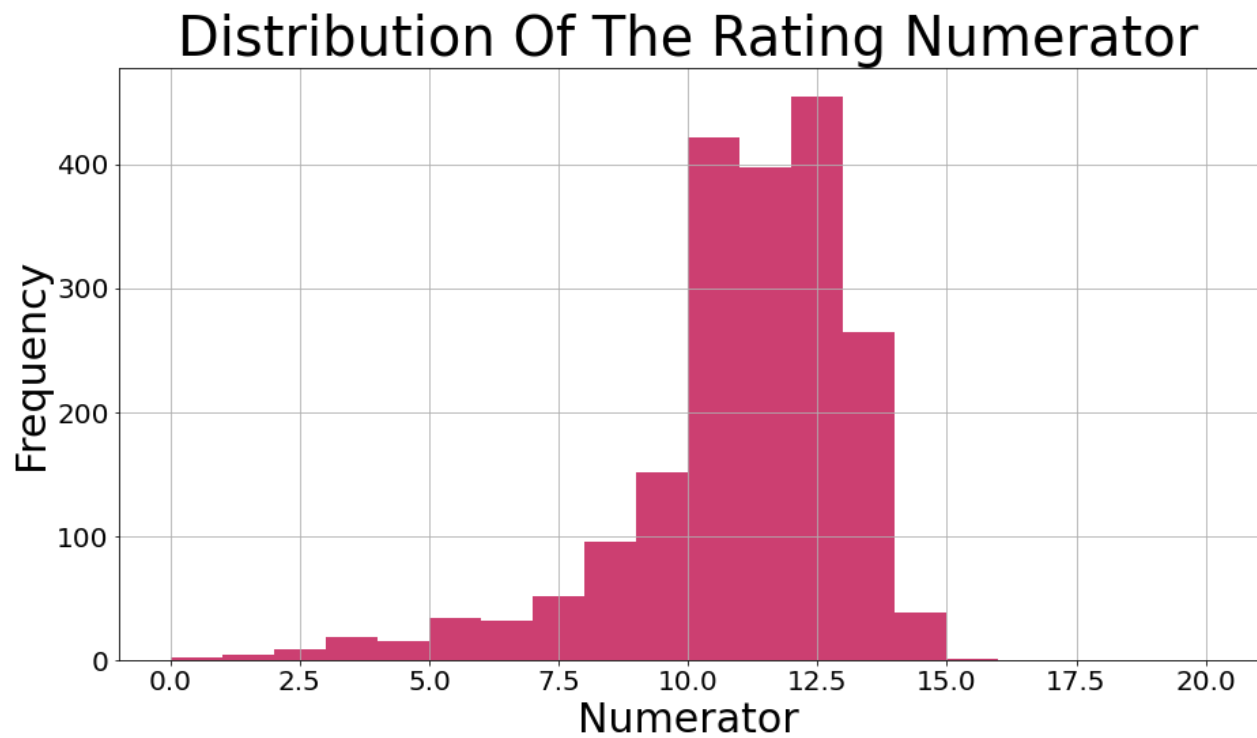
About the Data

To analyze the tweets from WeRateDogs, we have used three different source. The first source is an archive of the past tweets from @dog_rates provided via a CSV from Udacity. The second source is from the Twitter API used to retrieve more information about the tweets like number of each was retweeted. The third data source provides us the predicted dog breed in each tweet's image programatically determined from a neural network. This was also provided by Udacity. This third source is particularly important as it was not determined how accurate the predictions were.

QUESTIONS

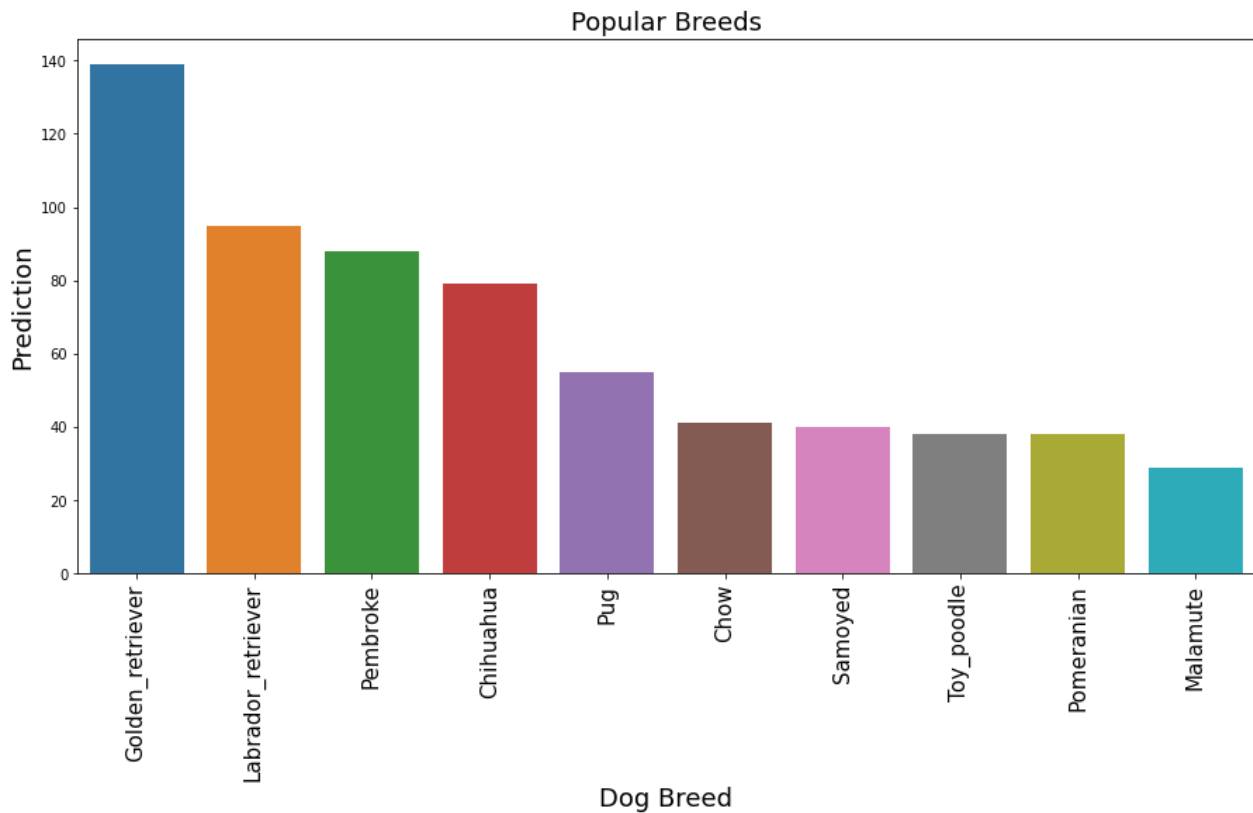
- Where the majority?
 - Most Popular Dog Breeds?
 - Retweet Count, Favorite Count and Ratings?
-

Where the majority?



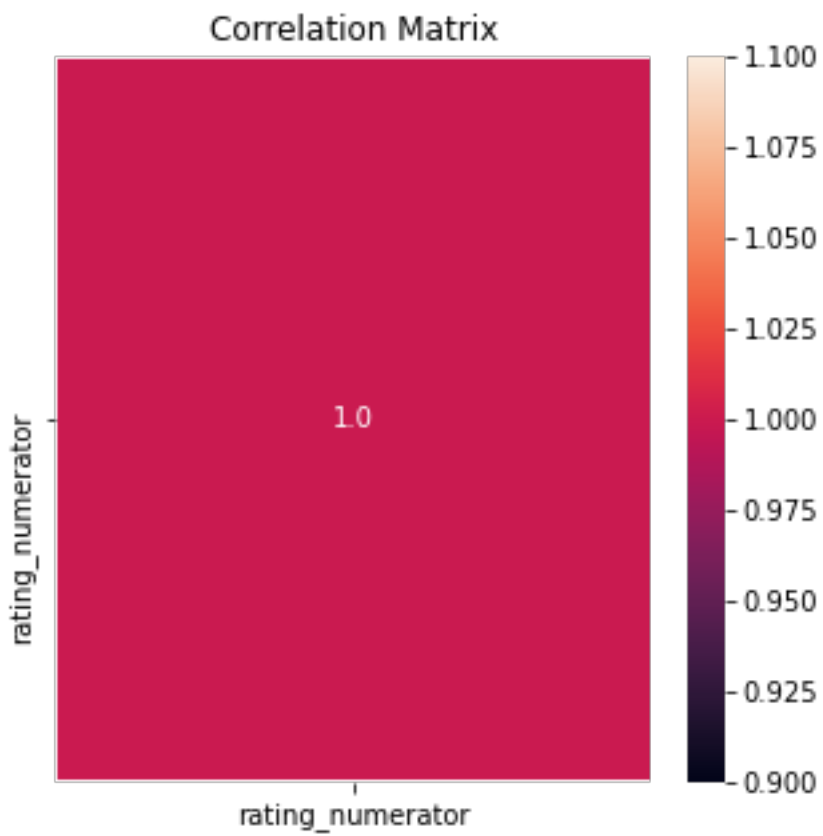
also notice that the most of majority of the numerators lies between 10 and 13.

Most Popular Dog Breeds?



As we see the after looking closely and analyzing the data on preference and retweeting, we clearly noticed that Most Popular Dog Breeds in terms of followers' preference and the majority is Golden retriever with a 140 difference as it comes with a number in terms of preference Labrador Retriever with 90 ..

Retweet Count, Favorite Count and Ratings??



As it shown above, 1.0 Retweet Count Favorite Count and Ratings.