

Project Analysis

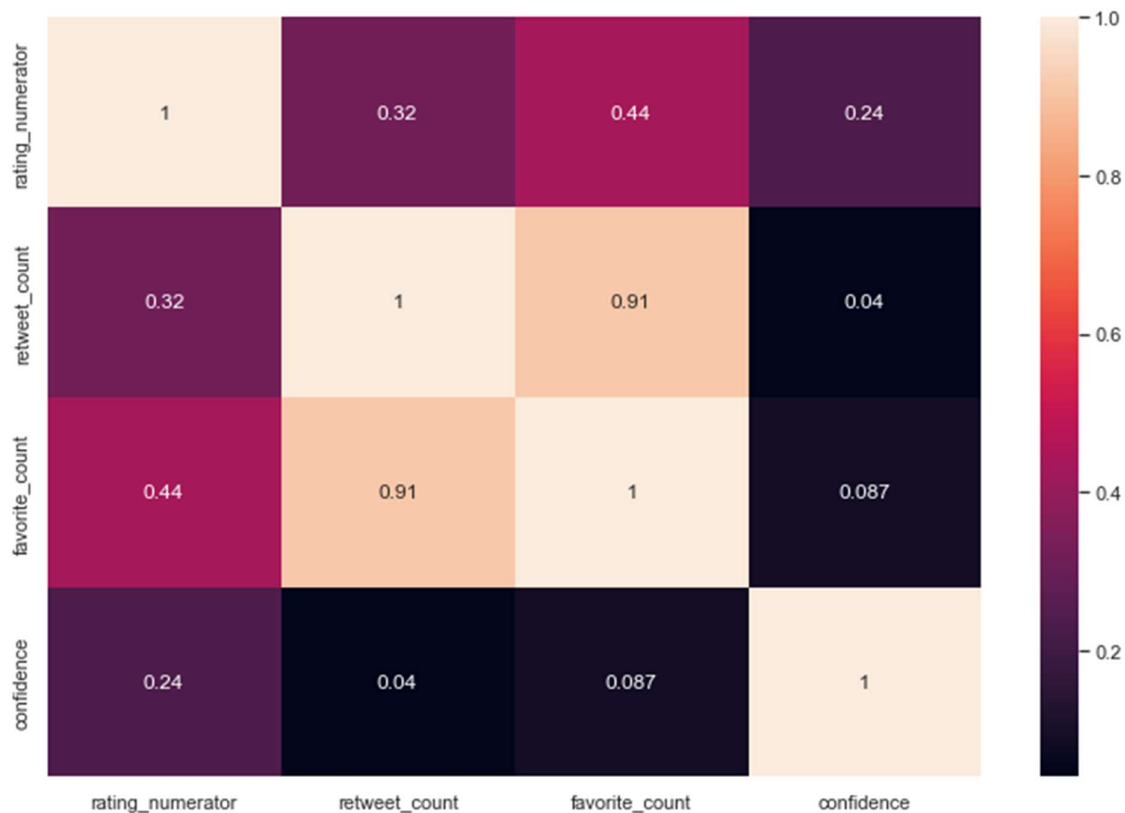
Introduction

The dataset used in this project was combined from a CSV document provided by WeRateDogs through Udacity, JSON text document provided by Udacity and a TSV, which was gotten from a Udacity on Udacity website.

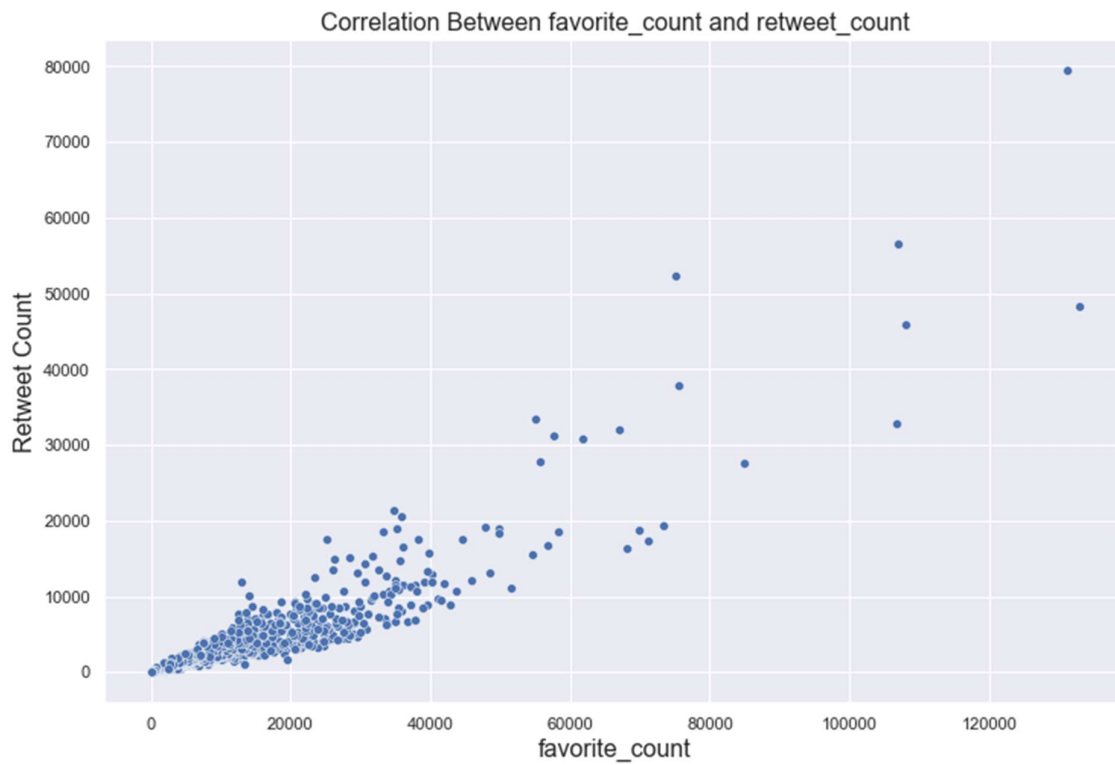
Data Analysis

After combining the datasets, it was trimmed down to 1906 rows and twelve columns and the analysis is carried out to answer the following questions;

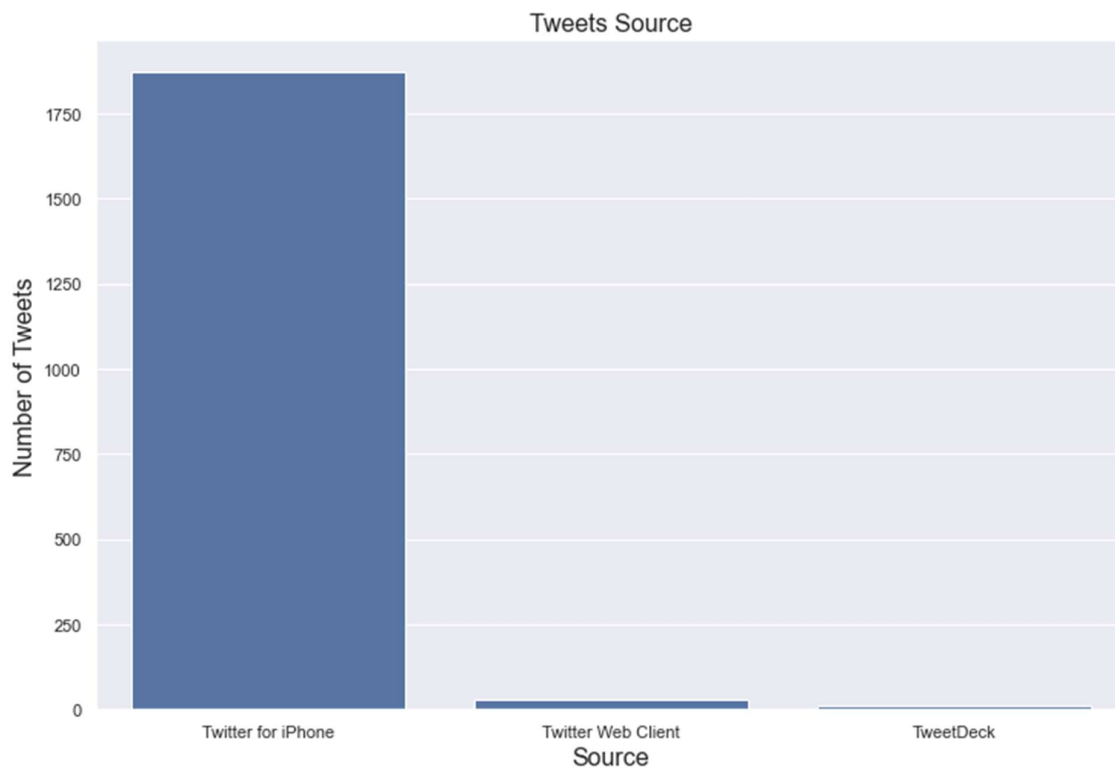
- What is the top source, users tweeted from?
- The top 10 most popular dog breed in the tweets
- Top rated dog based on retweets
- Top rated dog based on favorite count
- The winner of the contest, which is top rated based on dog rating and retweet count and favorite count



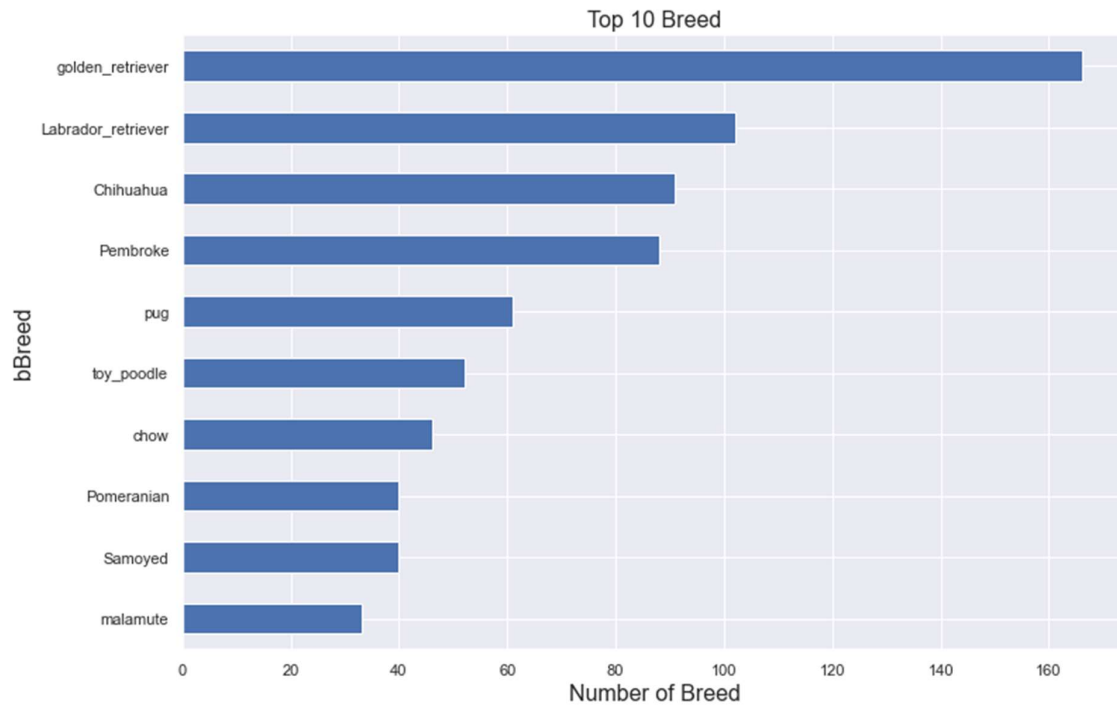
This plot shows a strong correlation between favorite counts and retweet count which strengthen some of the questions we already want to explore.



This plot further confirm the correlation between favorite count and retweet count showed by the heatmap



The top source used by users is the Iphone as you can see from the plot above.



This plot showed the top ten breed in the dataset, the top dog breed is Golden Retriever, followed by Labrador Retriever, after which the Chihuahua comes third in the ranking



This is the top rated retweeted dog image by users



This dog image is the top rated image based on favorite count



This is the image that won the contest and was chosen based on top dog rating, top retweet count and favorite count.

The project introduction did not mention any prize given, by a large it was a very interesting analysis, and I believe for the period of the contest, WeRateDogs handle will have be very popular.