

Analysis and Insights

By Tanmay Jain

Introduction

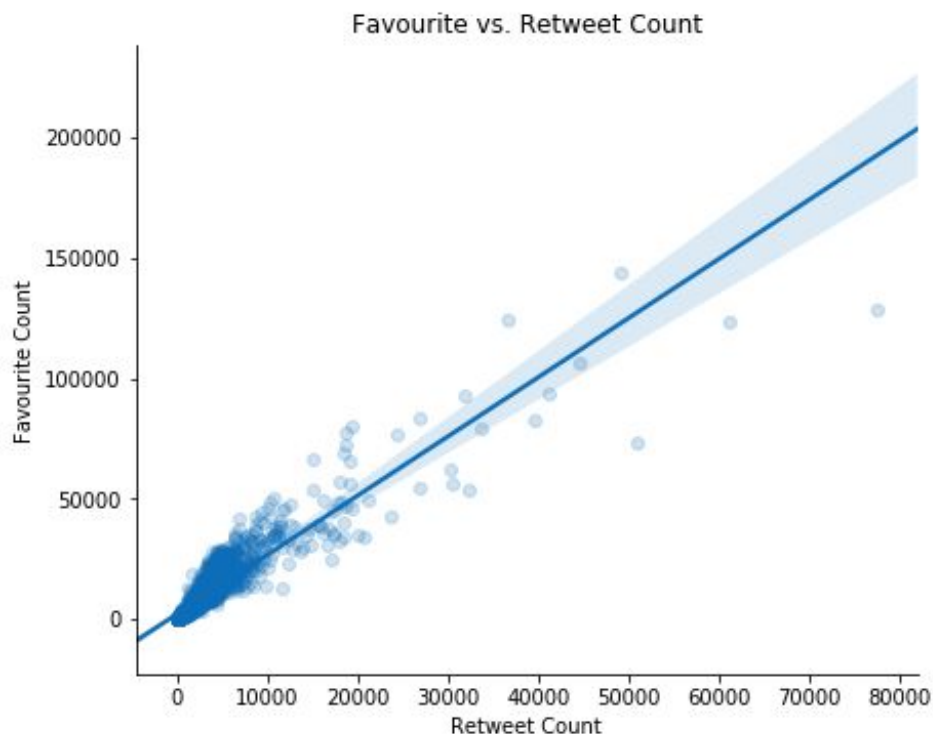
This report is a part of Udacity Data Analyst Nanodegree program. This project involves wrangling of data from various sources associated with tweets from the user

WeRateDogs(@dog_rates). WeRateDogs rates pictures of people's dogs. Instead of using the typical 10/10 format, it rates in a fun way and gives a rating more than 10 because they say Dogs always deserve a 10 and sometimes more.

This analysis report answers some question like "Which dog breed is the most common?" by analysing the data from the tweets.

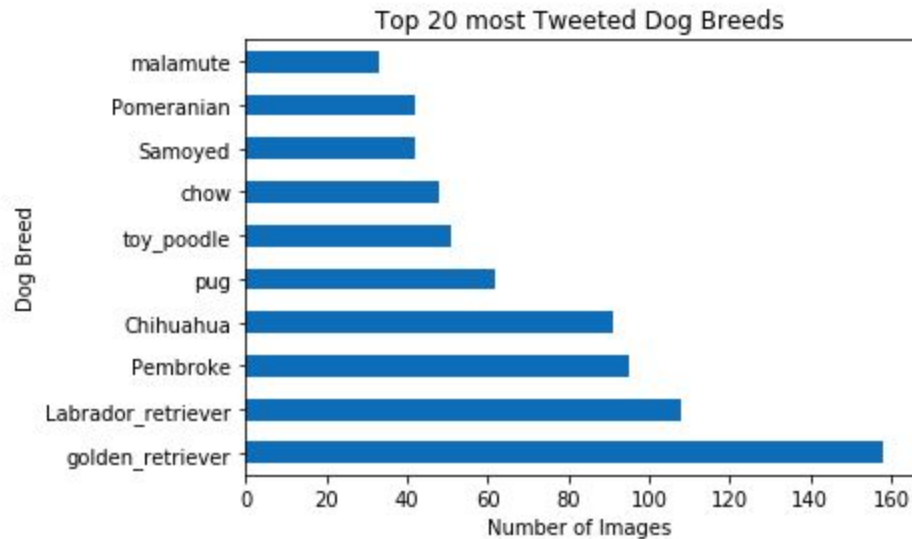
Favourite VS Retweet Count

In Figure 1, it can be seen that favourite and retweet counts are highly positively correlated. For about every 4 favourites, there is 1 retweet. The majority of the data falls below 50000 favourites and 20000 retweets.



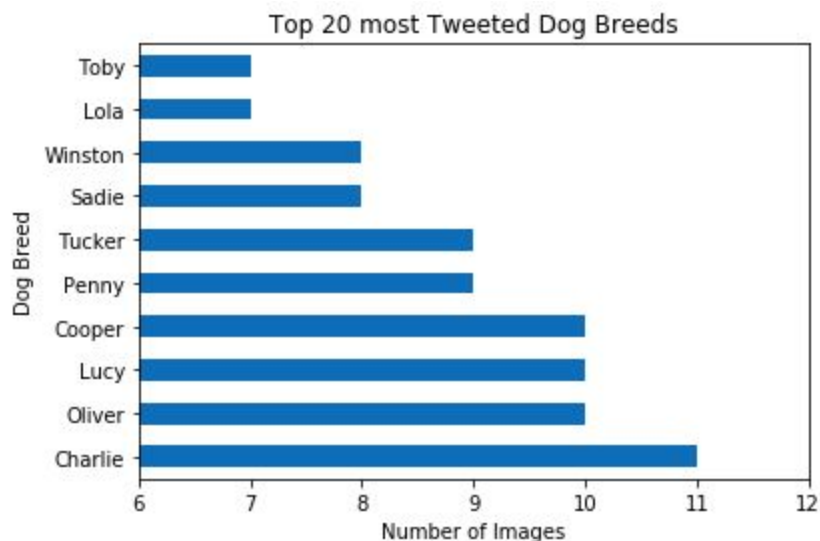
Most common Dog Type

On viewing the top 10 most tweeted dog breed, we can see that Golden Retriever is the most common dog breed tweeted by a huge margin, followed by Labrador and Pembroke.



Most Common Dog Name

On viewing the top 10 most common dog names, we conclude that Charlie is the most common dog name, followed by Oliver, Lucy and Cooper.



STANDARDIZED RATING OVER TIME

To standardize the ratings, I calculated a value of numerator divided by denominator. I was most curious to see if over time, as the account became more popular and people associated the above 10/10 ratings with being funny, that the higher ratings would become more prevalent. Indeed, as shown, it appears that over time the frequency of ratings below 1 decreases.

