Data wrangling WeRateDogs project

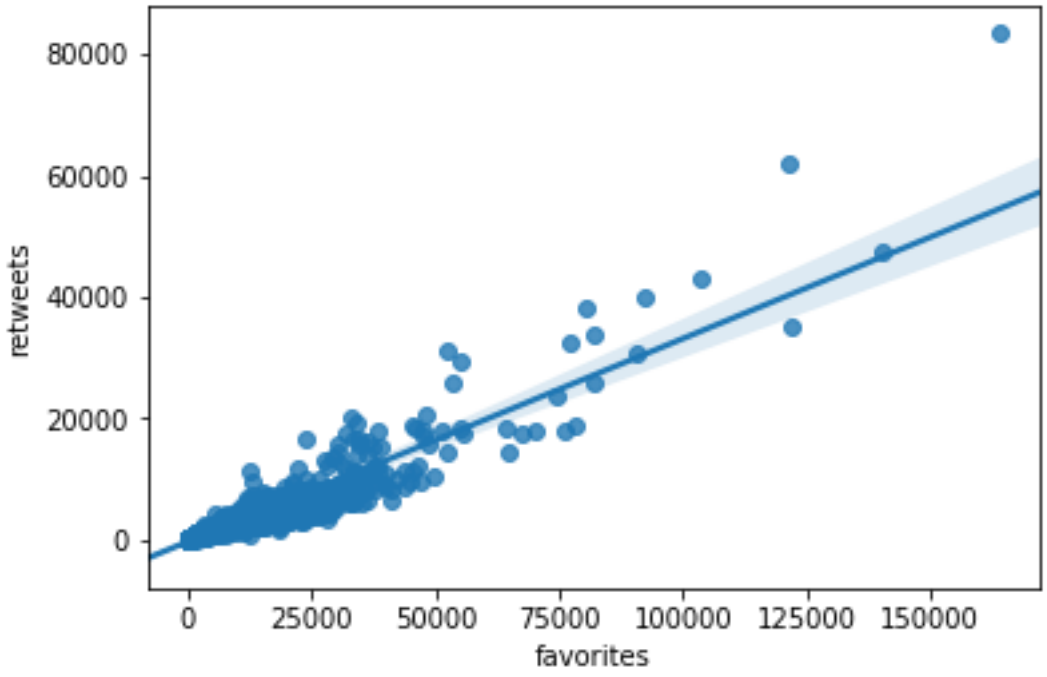
Analysis report

Introduction

This report was written as part of the Udacity nanodegree 'Data wrangling WeRateDogs' proj ect. It is a basic analysis of data extracted from the Twitter API from teitter account WeRate Dogs.

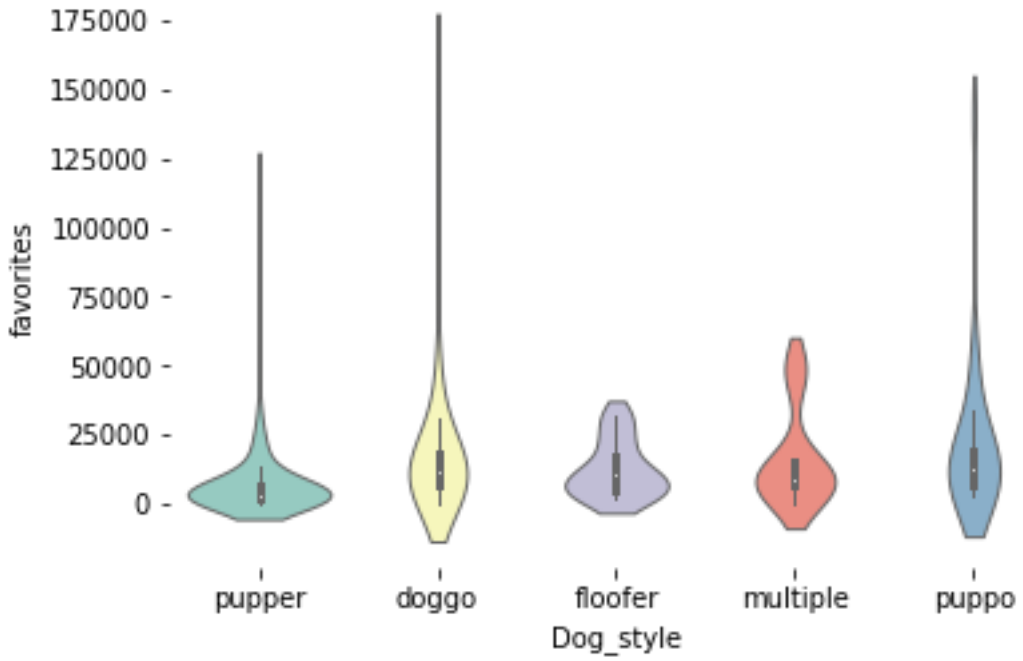
1.

At the begging I was curious on the corelettion between 'Rwteets' and "favprites" . From t he plot above, It is obvious that there is a strong correlation between the two values. Mor e retweets mean more favorite.



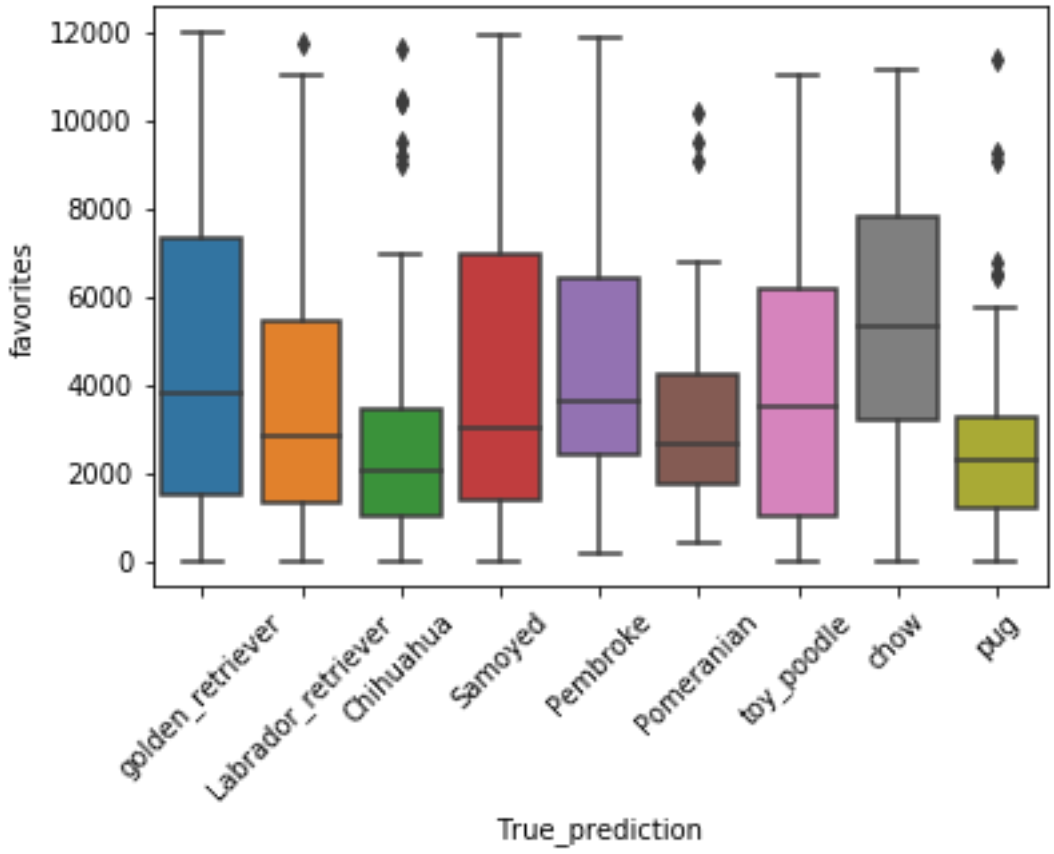
2.

Aftee that, I separate the dogs on 'Dog Syles' and look for the most popular and unpopular style. We can see that 'doggo' and 'puppo' styles collect more 'favorites'. On the other hand, pupper is closer to 0.



3.

Finally, I thought that It would be interesting to find out which breed is the most popular and witch the most umpopular. We can see that the most popular breed is the Chow and the most upopular is Chihuahua



4.

The most popular tweet is the one with tweet_id= 744234799360020481. It is a tweet with 163.982 favorites. At the picture there is a Labrador retriever doggo.

favorites	retweets	user_followers	user_favourites	Dog_name	Dog_style	True_prediction	c
163982.0	83426.0	7686882.0	140932.0	NaN	doggo	Labrador_retriever	

The same tweets is also the most retweeted.