

Analyzing and Visualizing WeRateDogs twitter archive data

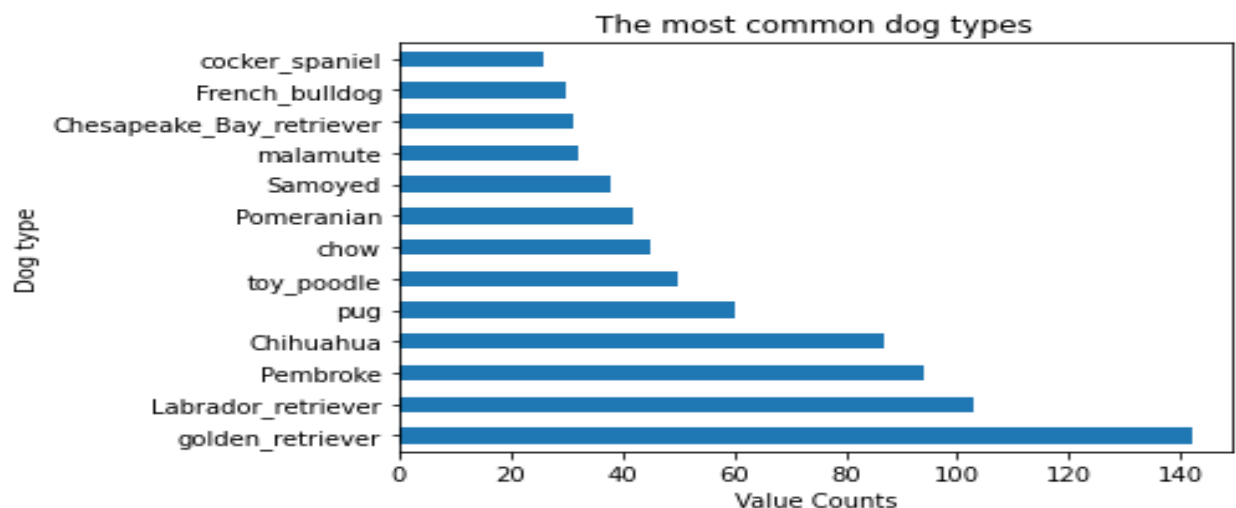
Introduction

First, WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators however are almost always greater than 10 (11/10, 12/10, 13/10, etc.). In theory the rates should be 1 to 10. However, WeRateDogs disagree, they admit almost all dogs deserve a 10 and sometimes more than that.

WeRateDogs has over 6 million followers and has received international media coverage. One of those coverage was about the quote “they’re good dogs Brent”, which was an exchange in which WeRateDogs shut down a person having an issue with its rating system in humorous ways.

In this project I will analyse the WeRateDogs twitter archive data and try answer some questions like: Which dog type (breed) is most common dog in the tweet dataset? What dog type has the highest average rating? Which dog type whose tweets related to it have the most favorites/likes? These questions and other insights:

Which is the most common dog type in the data set?



WeRateDogs has over 6000+ tweets. However, after cleaning I was able to analyze about 1500+ tweets, and in that data set the most rated dog was golden retriever with more than 140 ratings records.

Which dog type has the highest average rating?

Bouvier_des_Flandres has the highest average rating with average rating of about 13 out of 10. Below are some images.



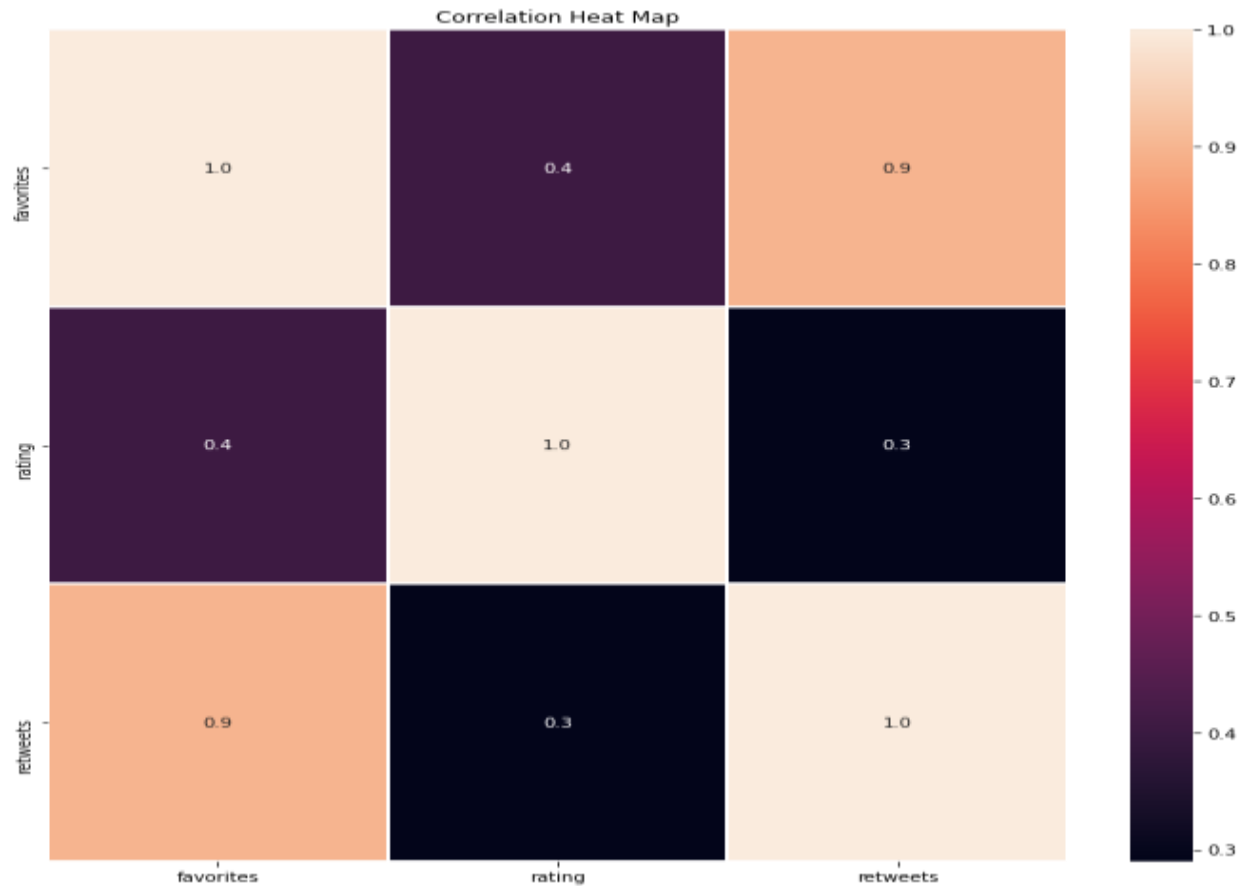
Bouvier_des_Flandres image source: google.com

Although I am not a dog person one should admit this creature looks real cute.

Which dog type whose tweets related to it have the most favorites/likes?

It seems tweets related to the golden retriever had the most likes about 1583506.0 Likes to be exact. However in this case we can't say that means people favored them, but it may be due to the fact that golden retrievers was the most common dog type thus more tweets were about it than the rest of the other types.

Does tweets with higher ratings get more likes or more retweets?



As shown in correlation heat map above dog's ratings is weakly correlated with the tweets retweets or likes. The correlation between dog rating and like is about 0.4 which is weak but positive correlation as shown in the top middle square. The correlation between ratings and retweets is even weaker but positive as shown in the bottom middle square only about 0.29 only.