

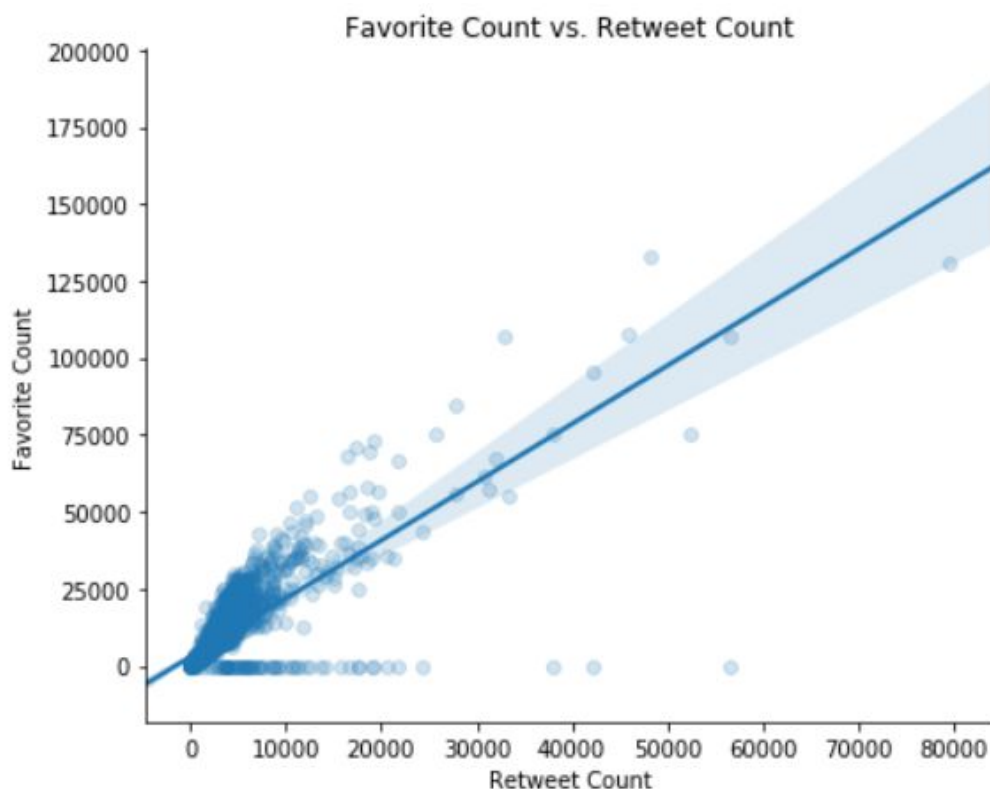
# WeRateDogs – Insights into the @dog\_rates Twitter page

## Introduction

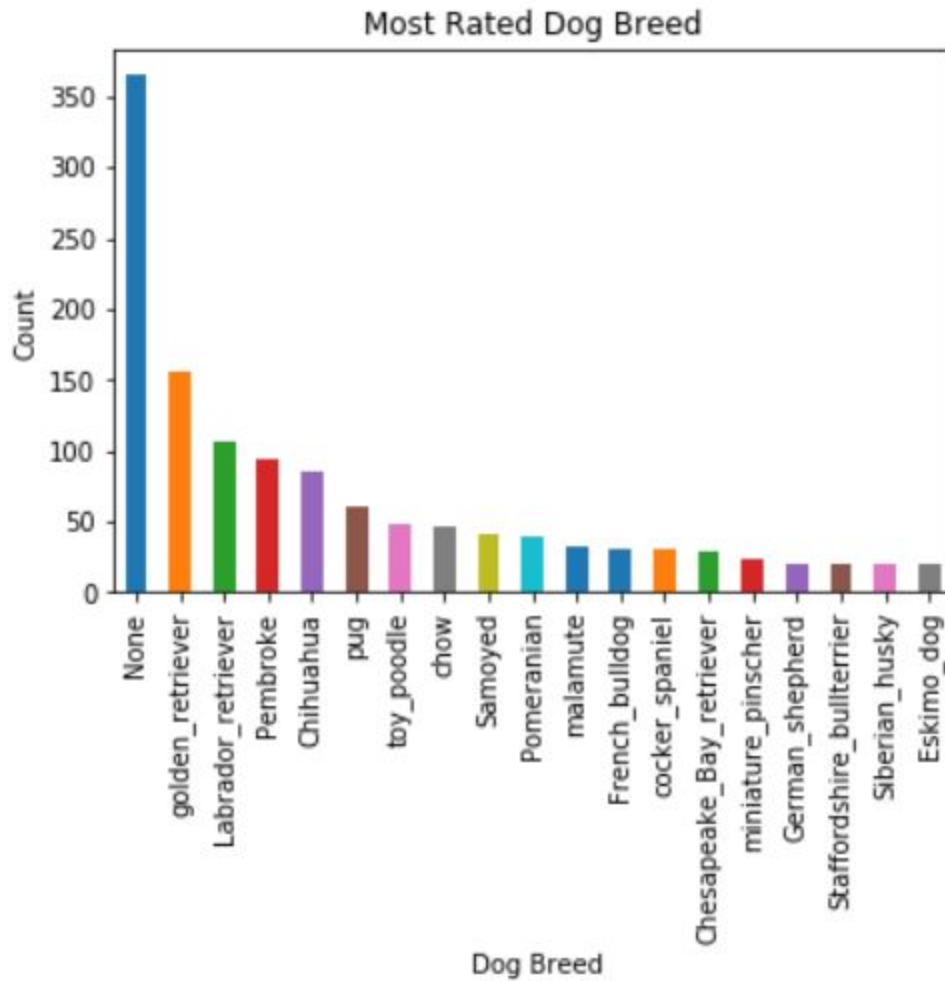
For this data wrangling project, the dataset wrangled here is the tweet archive of Twitter user @dog\_rates, also known as WeRateDogs. WeRateDogs is a Twitter account rating dogs of people with a witty remark about the dog.

Here in this project, I have gathered the data from three sources, then assessed it, and after careful visual and programmatic assessment cleaned it by removing some of the quality and tidiness issues, so as to make the dataset ready for some analyzations. Finally, I performed some visualizations so as to infer interesting facts about the project.

Visualisations :

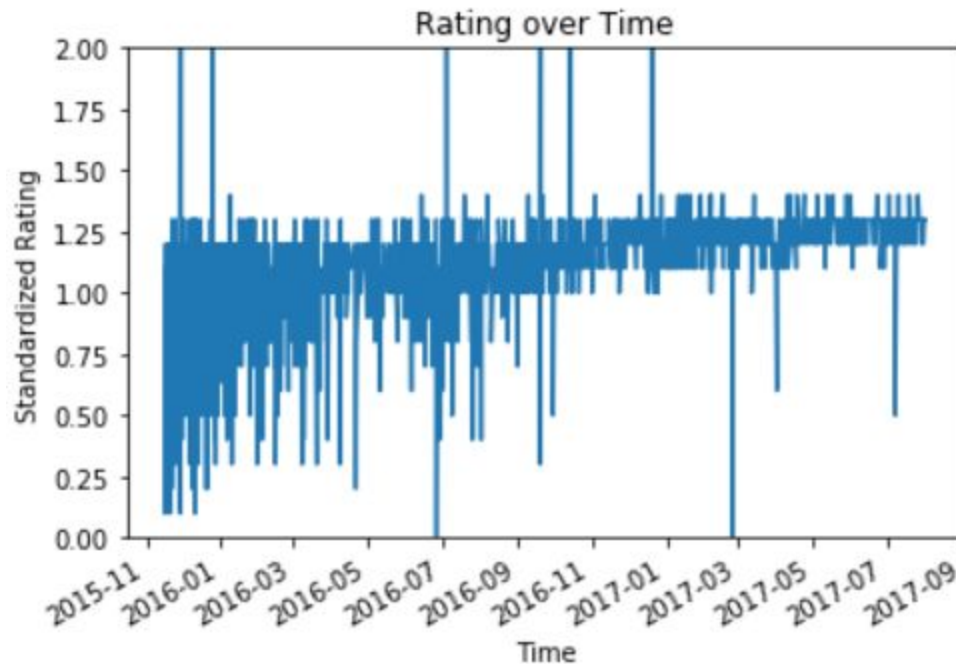


We can see that there is a positive correlation in the above plot between favourite and retweet count. We can say that the tweets being retweeted most are the favourite.



We can see that Golden retriever is the most common dog breed followed by the labrador retriever, with the eskimo dog being the rarest of all dog breeds.

Assumption : I have ignored the none rows where the dog breed is not provided.



The rating over time plot clearly shows the ratings through the years 2015-2017. Most of the rating are below 1 before the 2016-19.

Conclusion :

I have analysed and visualised three parameters after cleaning the data. Visualisations are a great way to understand more about the data faster, as it gives a clearer picture of the data. Thus, data visualization is a core part of data analysis which follows the data wrangling process.