

# Data Analysis

*Blogpost by Pankaj  
Jun, 2020*

## Introduction

This is the analysis of the tweets of the WeRateDogs twitter account. It has whopping 8.7 million Followers around the globe. So, @dog\_rates do is that, they rate people's dogs with a humorous comment about the dog. Some examples are as below,

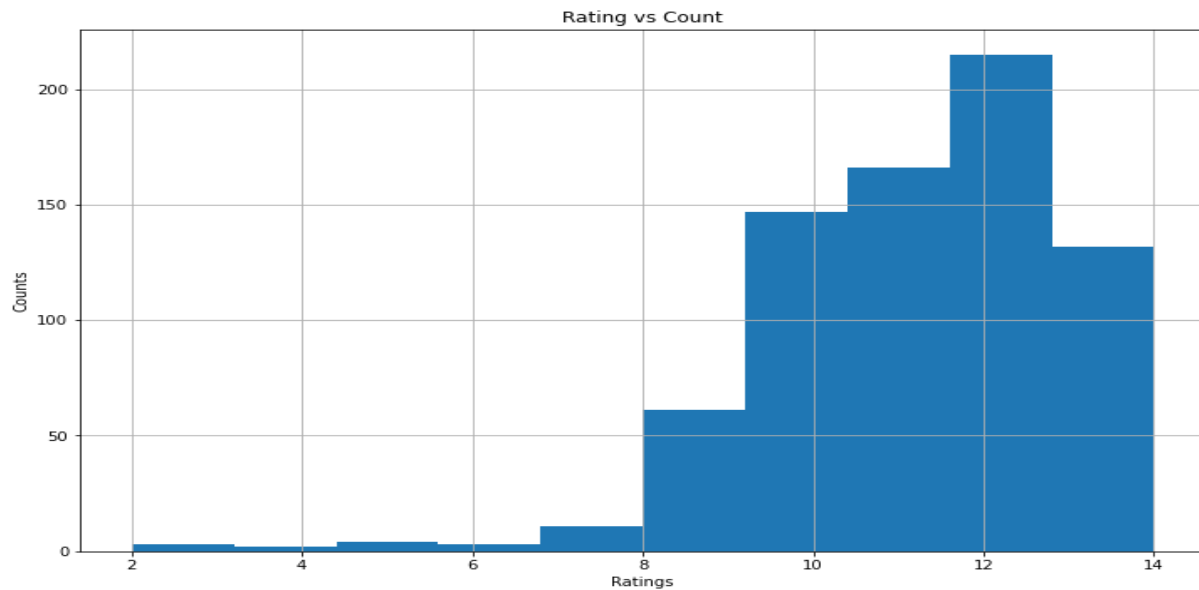


So, while scrolling through the tweets, some questions aroused in my mind like which dog got the most ratings or which breed of dog were mostly tweeted about from this twitter account, what were the frequent ratings. So to get the answer to it, let's switch to programmatically analysis of the tweets of the WeRateDogs twitter account.

## Analysis

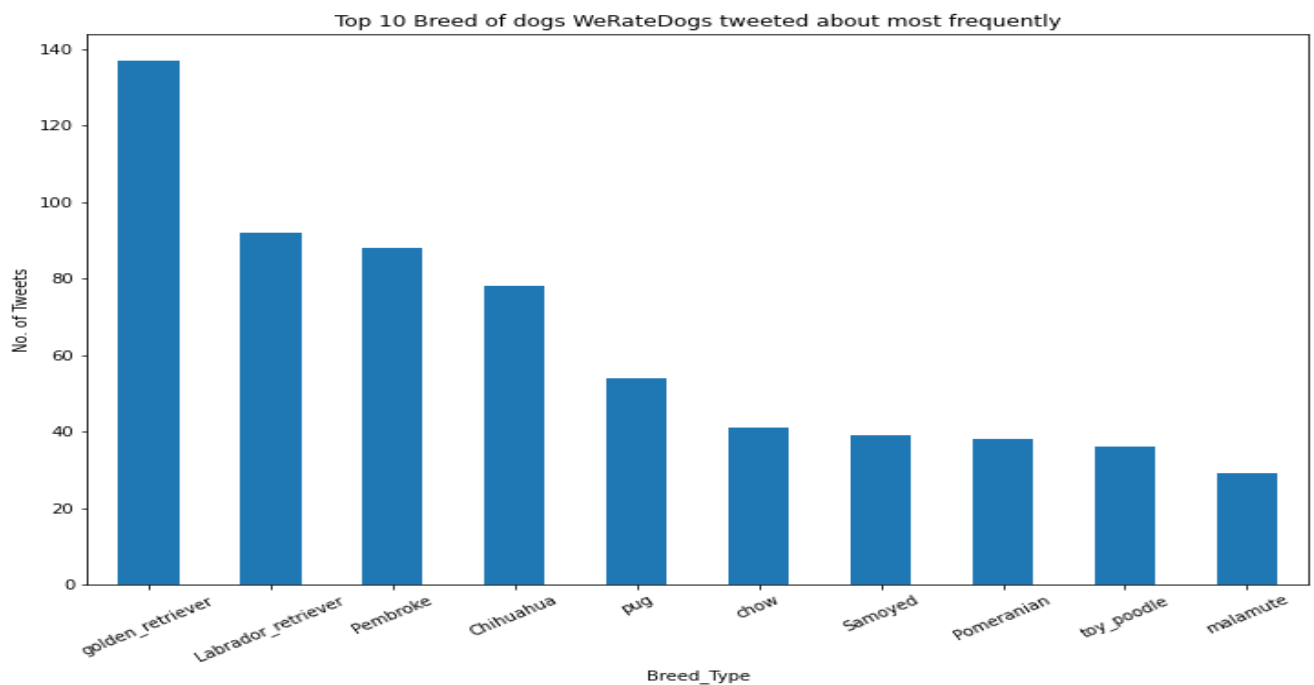
The WeRateDogs Twitter archive contains basic tweet data for all 5000+ of their tweets. So, let's start the analysis part by first question,

a.) What were the most frequent rating among the tweets?



**Insight:** - Most of the ratings given by WeRateDogs in tweets are ranging from 9-14.

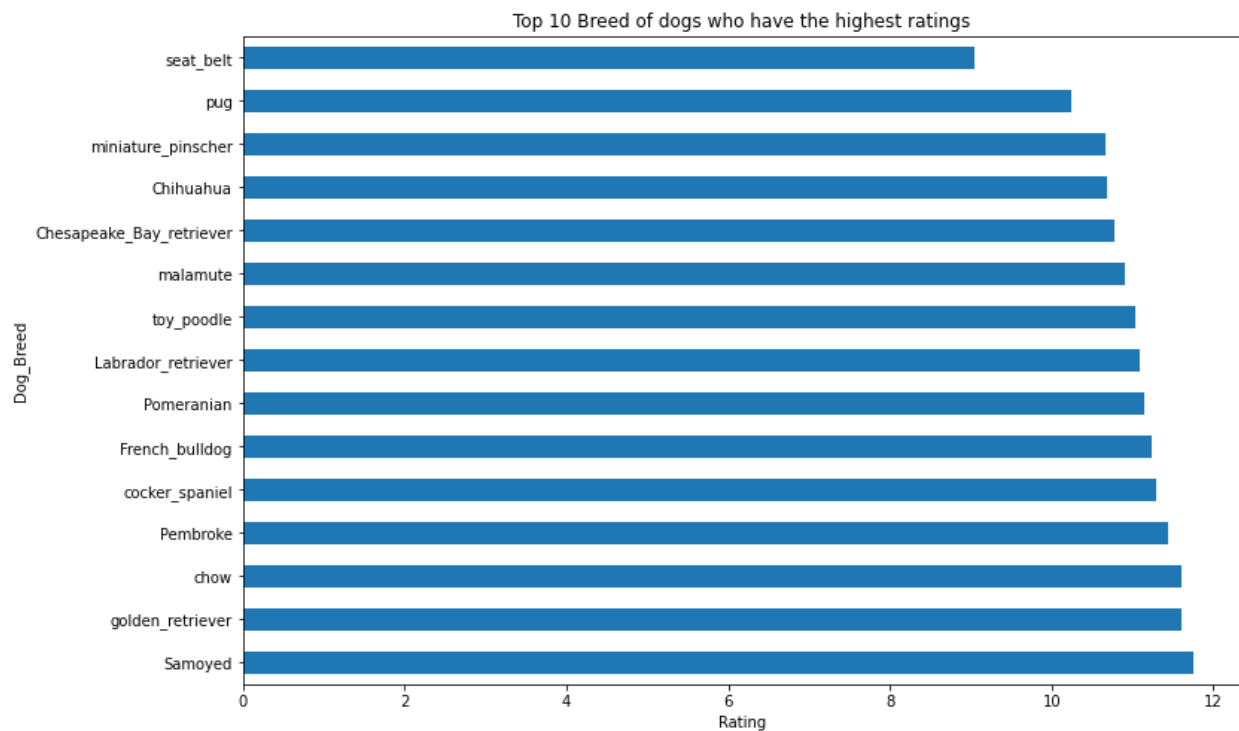
b.) Which breeds of dogs WeRateDogs most frequently tweeted about?



**Insight:** -The top 5 breed types who were most popular were **Golden Retriever, Labrador retriever, Pembroke, Chihuahua and Pug.**

So it was clear from the graph that Golden Retriever is quite famous among the WeRateDogs tweets. Lets find out if he had some good ratings as well or not.

c.) Finding the top breeds of dogs which had the highest average rating

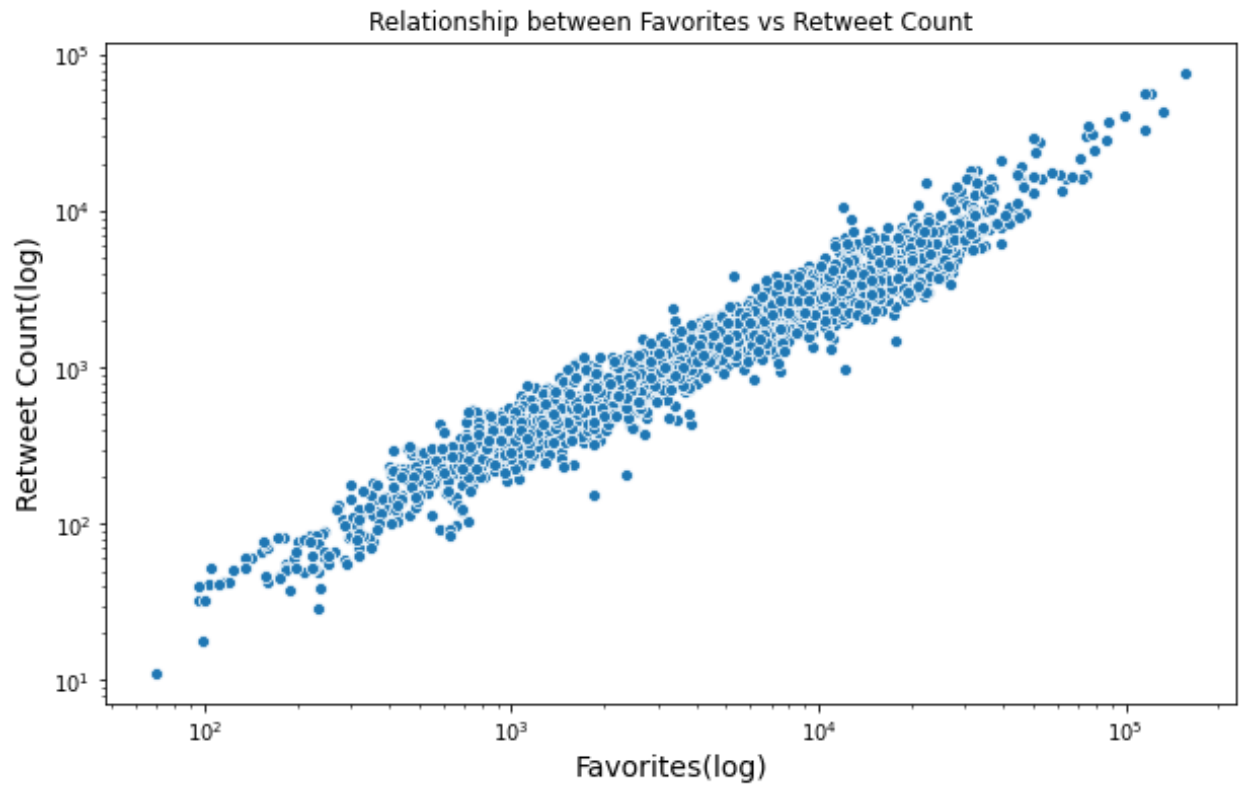


So, from the graph it is clearly seen that the Samoyed, Chow, Golden retriever were among the top breeds who had the best ratings relative to all other breeds.

In this I took the breeds which had at least 20 tweets about them as it will be more accurate to take a larger data to find the average ratings.

While I was playing around with the data and getting to know more about the tweets and text that were posted and variables which were factored most etc I found a relationship.

d.) Relationship between Favorites vs Retweet Count



There is a strong correlation between the favorites count and retweet count i.e. if a person likes a tweet so he will most probably retweet it. Its looks reasonable as well .