

# Act Report from #WeRateADog analyze

SEBASTIEN HANICOTTE

MAY 2020

## Introduction

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. For this project, an archive containing more than 2.5K Tweets with lot of information was given (tweet\_id, votes, dog stage, dog name, source).

I used Twitter API to retrieve complementary datas for all those tweets like “Nb of Favourites”, “Nb of Retweets”.

I used also the Prediction Image TSV to enhance the datas with “Dog Race” and “Confidence”.

## Data Wrangling

Before working with datas and trying to display and analyze all those data, I had to clean all the data and perform some transformations to be more handy.

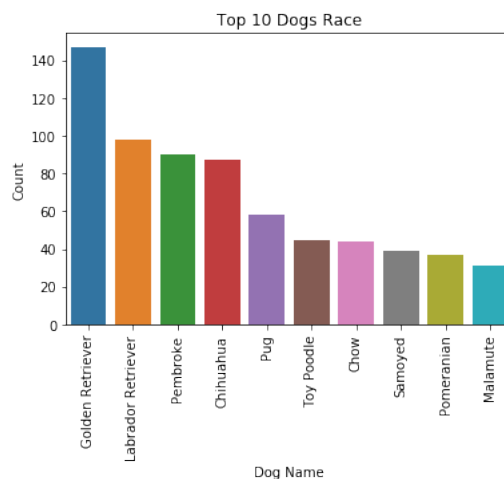
After all those cleaning stage, the remaining tweets to analyze were still above 1500.

## Insights

### Distribution of Dog Breed

From all those tweets, top 10 dog breeds are by decreasing order :

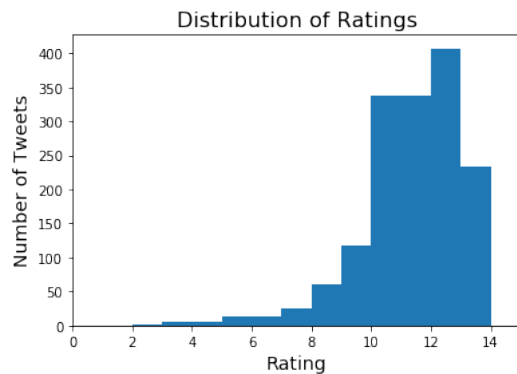
- Golden Retriever
- Labrador Retriever
- Pembroke
- Chihuahua
- Pug
- Toy Poodle
- Chow
- Samoyed
- Pomeranian
- Malamute



### Distribution of Ratings

Only ratings with a denominator of 10 were maintained in the dataset. But most ratings are between 10 and 14.

This could be explained by the fact that people like a lot those dogs and therefore they are giving them “huge” rate.



### A Strong relationship between Favourites and Retweets

There is a strong relationship between Nb of Favourites and Nb of Retweets. Does one criteria impact the other? One sure thing is that both are strongly related. By the way, it seems that globally there is higher nb of favourites than retweets. Perhaps that's because people are too shy to cry out their fondness for a specific tweet by sharing it to their followers but prefer to show the tweet owner that they love it.

