We Rate Doge – Project 4

Introduction

The dataset that we will be wrangling (and analyzing and visualizing) is the tweet archive of Twitter user @dog_rates, also known as WeRateDogs. 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, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage. WeRateDogs downloaded their Twitter archive and sent it to Udacity to use in this project. This archive contains basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017.

Goals

Main tasks in this project are as follows:

1. Data wrangling, which consists of:

- Gathering Data
- Assessing Data
- Cleaning Data

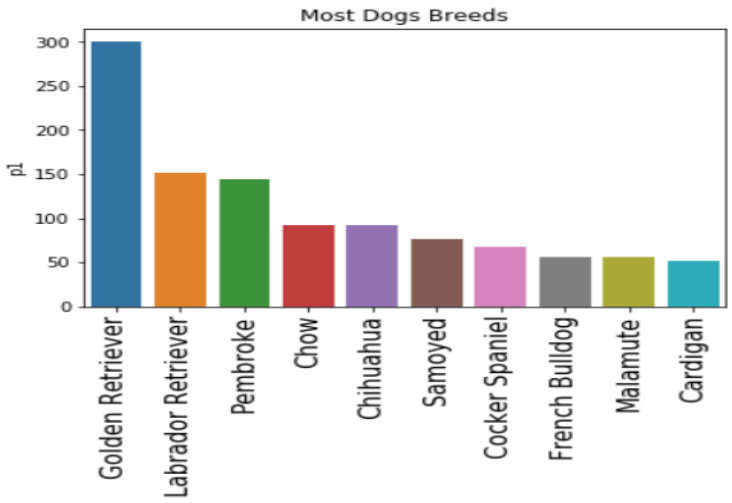
2. Storing, analyzing, and visualizing your wrangled data

3. Reporting

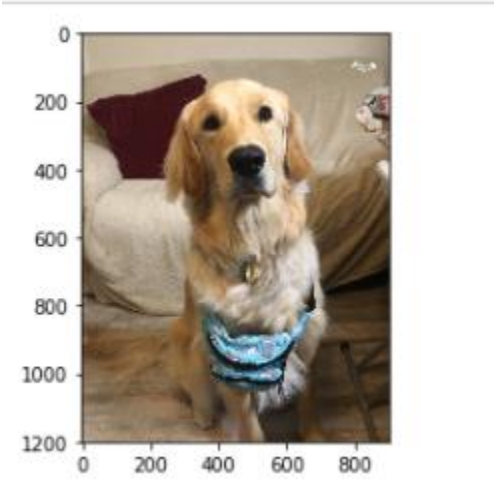
General

If we are look to the data, we get more understanding about WeRateDogs data, let's get started.

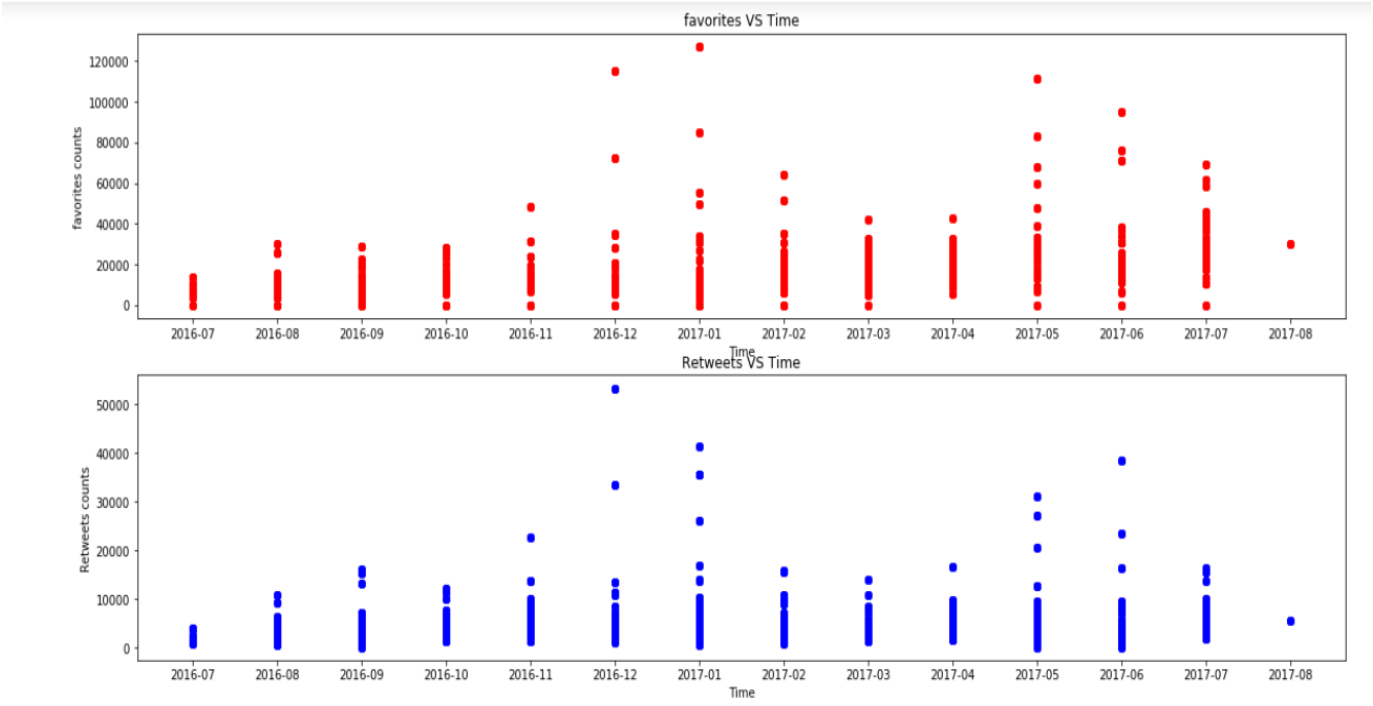
1. When we're are the looking for the most popular breeds as with high value prediction (p1), we found the "Golden Retriever" has most counts



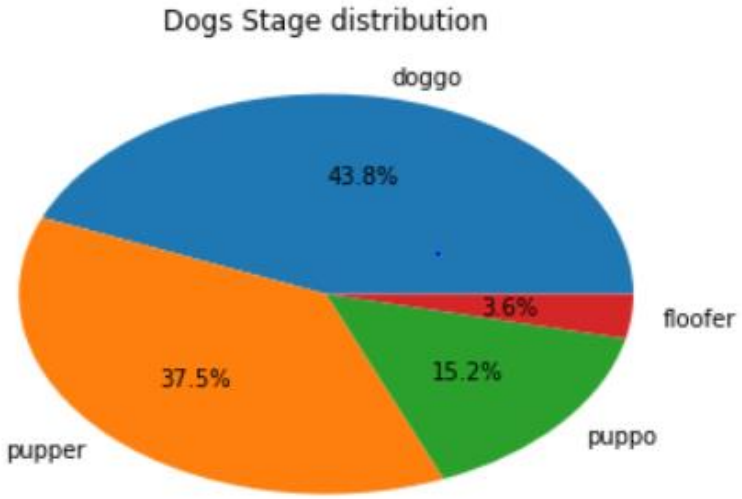
2. Picture of the "Golden Retriever", if you don't know this kind of dog



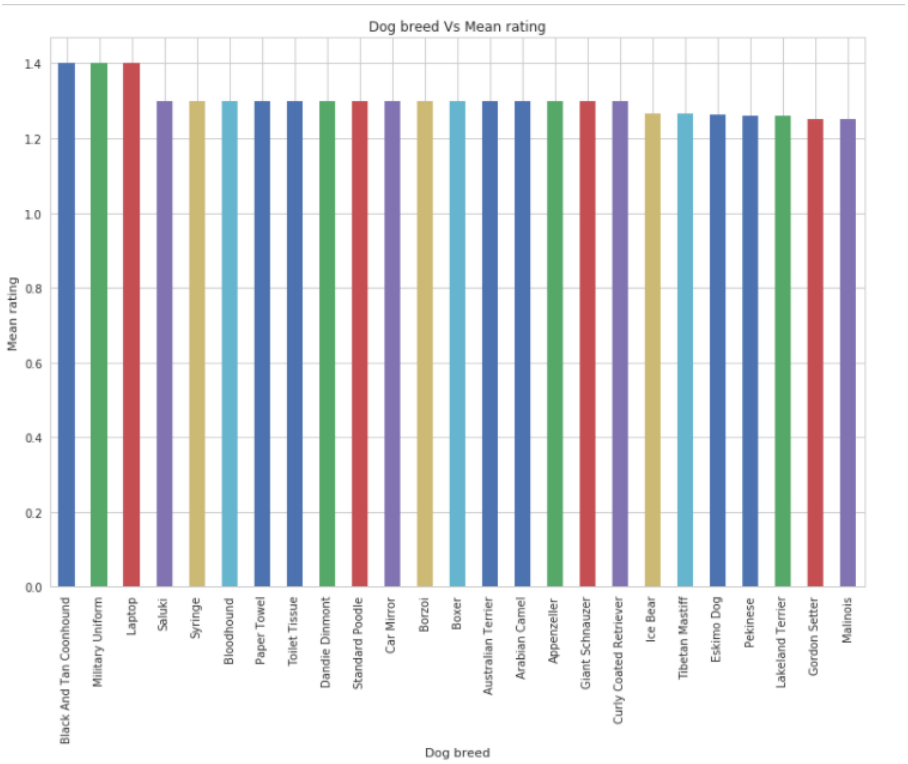
3. We look to the retweets and favorites counts through Time, in 2017 – 01 and 2017 – 07 have the most counts "remove the outliers".



4. We found that the most dogs depend on the data are in doggo stage.



5. The dog breed has ahigh rating is “Black And Tan Coonhound” with 1.4 rating



Conclusion

After these simple points, we reach to these points:

1. From above visualization we are looking for the most dog breed population
2. We are looking for the rating of each dog breeds.
3. We look the most popular stages.
4. The count of favorites and retweets during the time