Act Report

The clean dataframe is a result of a chained process of gathering, assessing, cleaning and visualyzing data from WeRateDogs twitter as a proof of stake for Data Wrangling module (Udacity, july 2019).

Dataframe1 was composed by 1887 tweets and all the intrinsic data such as: url, timestamp, dog class, url and rating. As long as dataframe2 was composed by 2334 retweets/favs posts and data from image predictor algorithm.

Figure 1 shows the rating distribution among numerator ratings in dataframe master. The ratings are between near zero to 1.75, being 1.25 the most common rating at all.
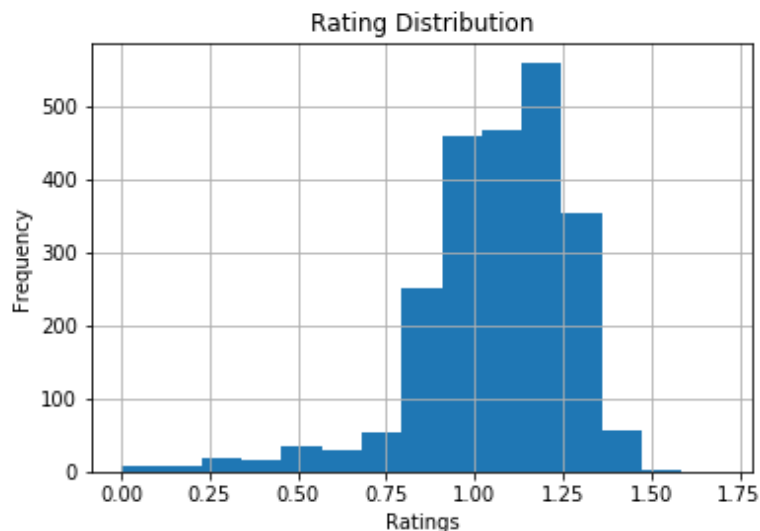

Figure 1 - Rating distribution.

Figure 2 shows the dog class distribution among doggo, pupper, puppo and floofer. These are specific to WeRateDogs so that most of rows has a null value for this variable. Pupper being the most common class by choice.

Figure 3 shows the description of dataframe 2 in terms of favorites and retweets. It's possible to check that were 2334 rows analyzed. There were a mean of 7794 favorited tweets and a mean of 2839 retweets.

By analyzing the image algorithm's reponse we can conclude that Golden Retriever was the most common dog predicted overall in WeRateDog twitter.
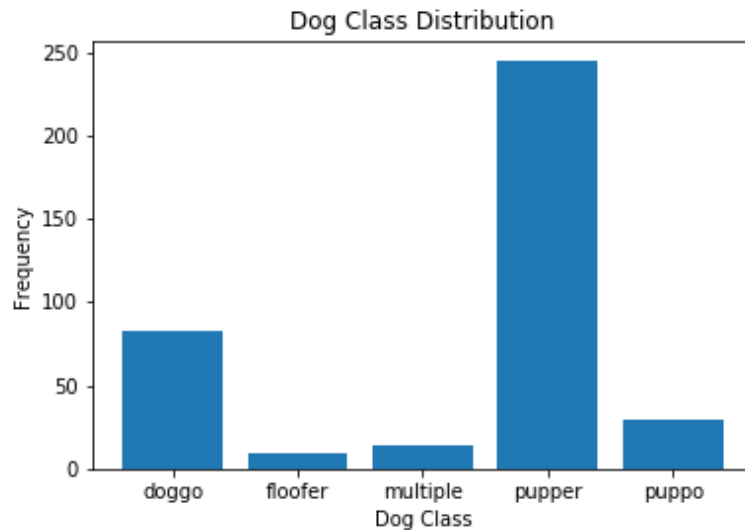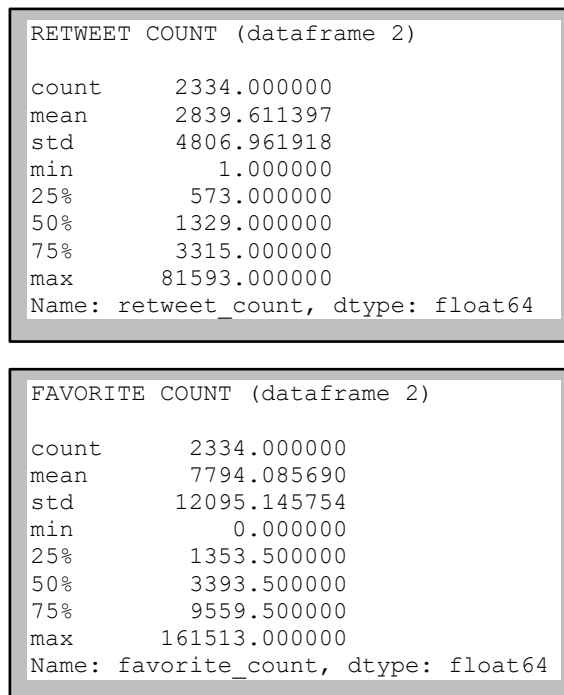
Figure 2 - Dog class ratings.

```
RETWEET COUNT (dataframe 2)

count      2334.000000
mean       2839.611397
std        4806.961918
min           1.000000
25%         573.000000
50%        1329.000000
75%        3315.000000
max       81593.000000
Name: retweet_count, dtype: float64
```

```
FAVORITE COUNT (dataframe 2)

count      2334.000000
mean       7794.085690
std       12095.145754
min           0.000000
25%        1353.500000
50%        3393.500000
75%        9559.500000
max      161513.000000
Name: favorite_count, dtype: float64
```

Figure 3 – Mean and median of retweets and favorites.

Final remarks:

1. The most frequent rating is 1.25.

2. The most common dog class is Pupper.

3. The most popular dog, based on image predictions, is a Golden Retriever.

4. The median and mean favorite count is 3393 and 7794 respectively

5. The median and mean retweet count is 1329 and 2839 respectively