Report: Act Report

My name is Adedayo Adediran. I am currently enrolled in the Udacity ALX-T Nanodegree Scholarship Program - The Data Analyst Track. For my second project, I was requested to seek insights into 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. WeRateDogs has over 4 million followers and boasts of considerable international media coverage. The Twitter Archive contains over 5,000 tweet data which was made available for analysis by WeRateDogs via Udacity.

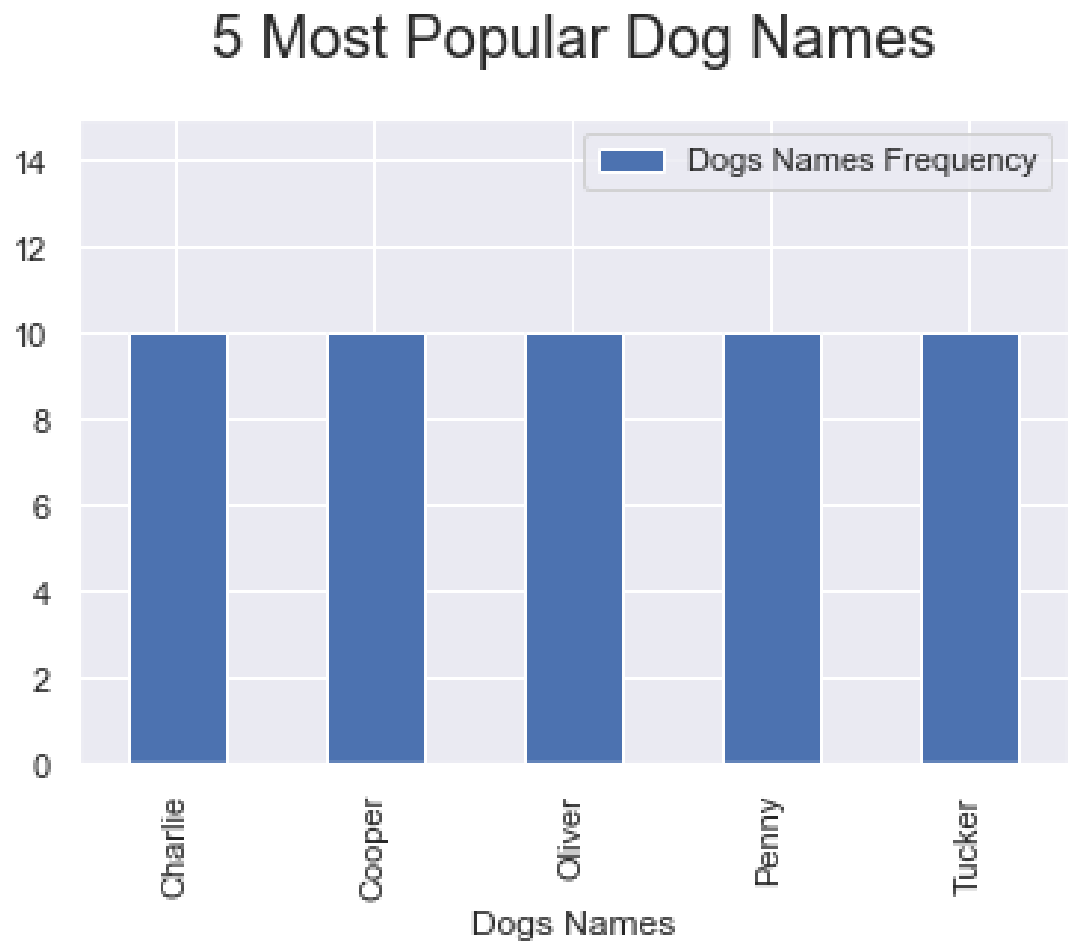
To assist in getting quality analysis and visualizations, I was obliged to gather important omissions in the Enhanced WeRateDogs Twitter archive directly from the Twitter API and import an already provided image prediction file. After importing the three files, I assessed them visually and programmatically for quality and tidyness issues using the right pandas syntax.

I proceeded to cleaning the dataset, which was important to getting an interesting and trustworthy analyses and visualization. This was particularly a tough nut to crack given the many issues with the WeRate dataset and the Twitter API data. At the end of the cleaning step which also included merging, I was left with 2,057 rows and 14 columns which I used for my analysis and visualization. I answered 10 questions using the cleaned and merged datasets.

Questions answered in my analysis:

1. What are the most popular dog names?

Answer: I looked at the most popular dog names in the master dataset and found that Tucker, Penny, Olivier, Cooper, and Charlie were the most popular dog names in the WeRate dog tweets. All occurred at least 10 times in the tweets. The visualization is shown below:



2. Which dog tweet was the most favorited?

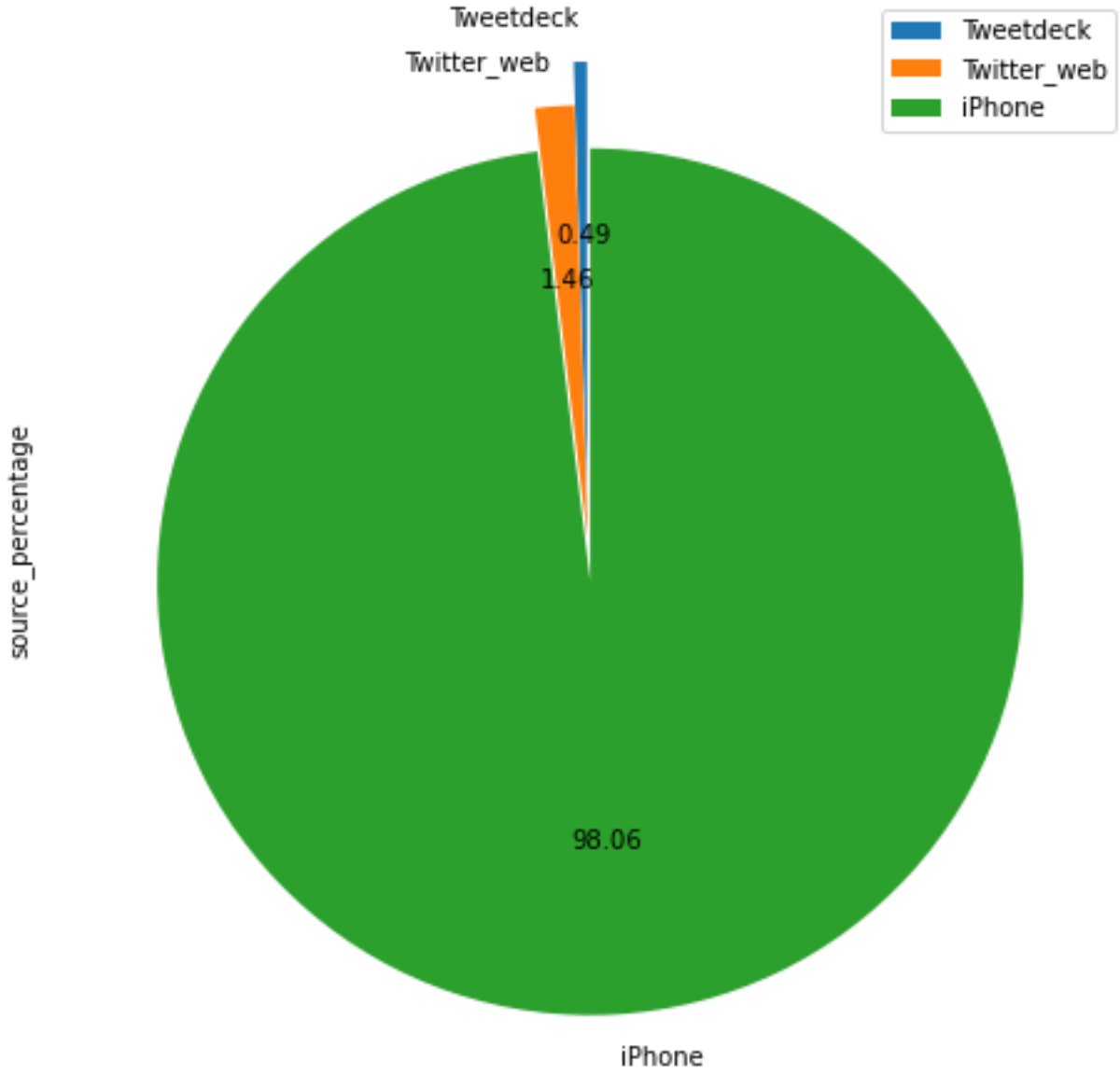
Answer: The most favorited dog tweet is the tweet with the tweet_id 744234799360020481. It was posted on the 18th of June 2016 at 18:26 pm. The dog stage is doggo, and the breed is Labrador retriever. It was favorited 144201 times.

3. Which dog tweet is the most retweeted?

Answer: The tweet with the tweet_id 744234799360020481 is also the most retweeted tweet with 70,316 retweets.

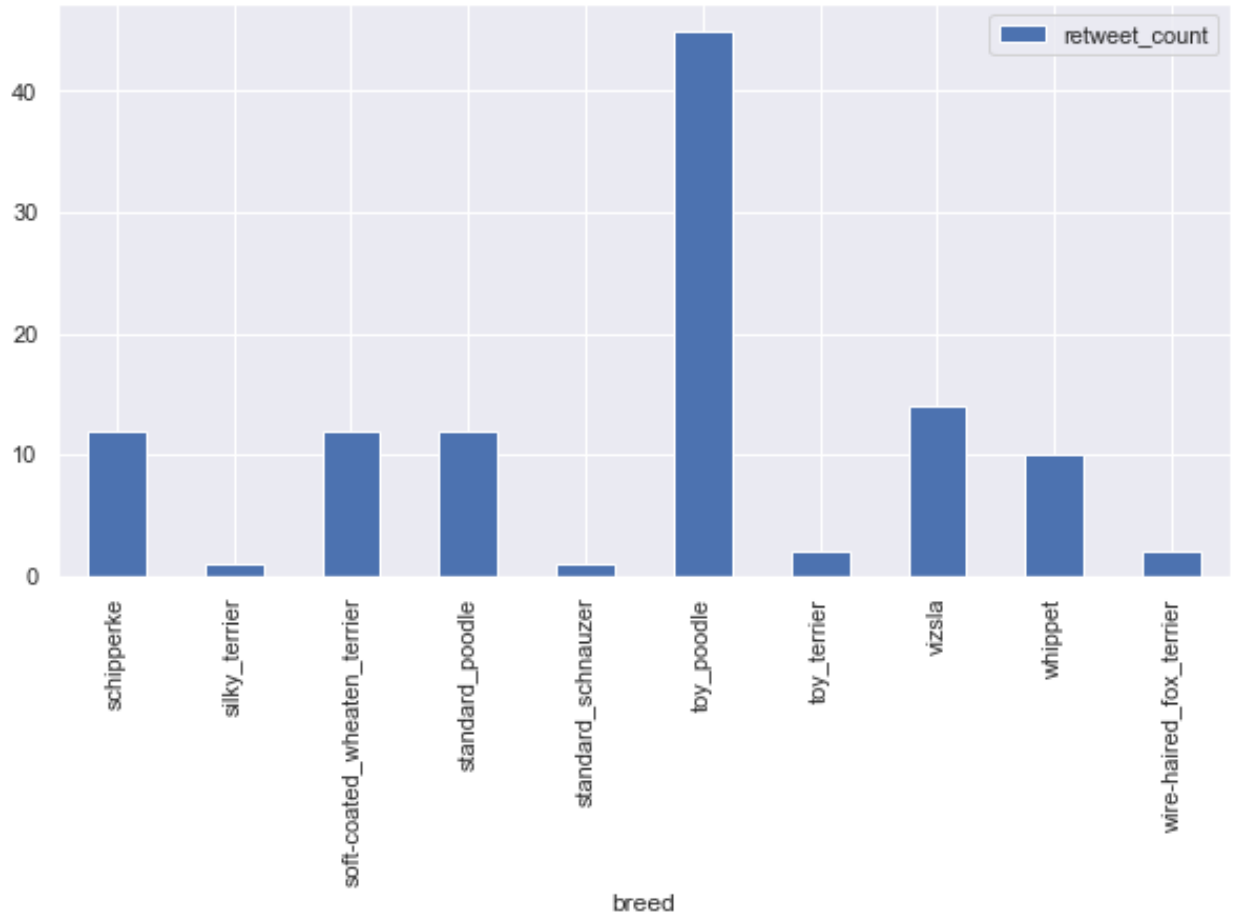
4. What are the most common sources of dog tweets?

Iphone was the most popular source of dog tweets. The visual representation of the sources of the tweet is provided in the pie chart below:



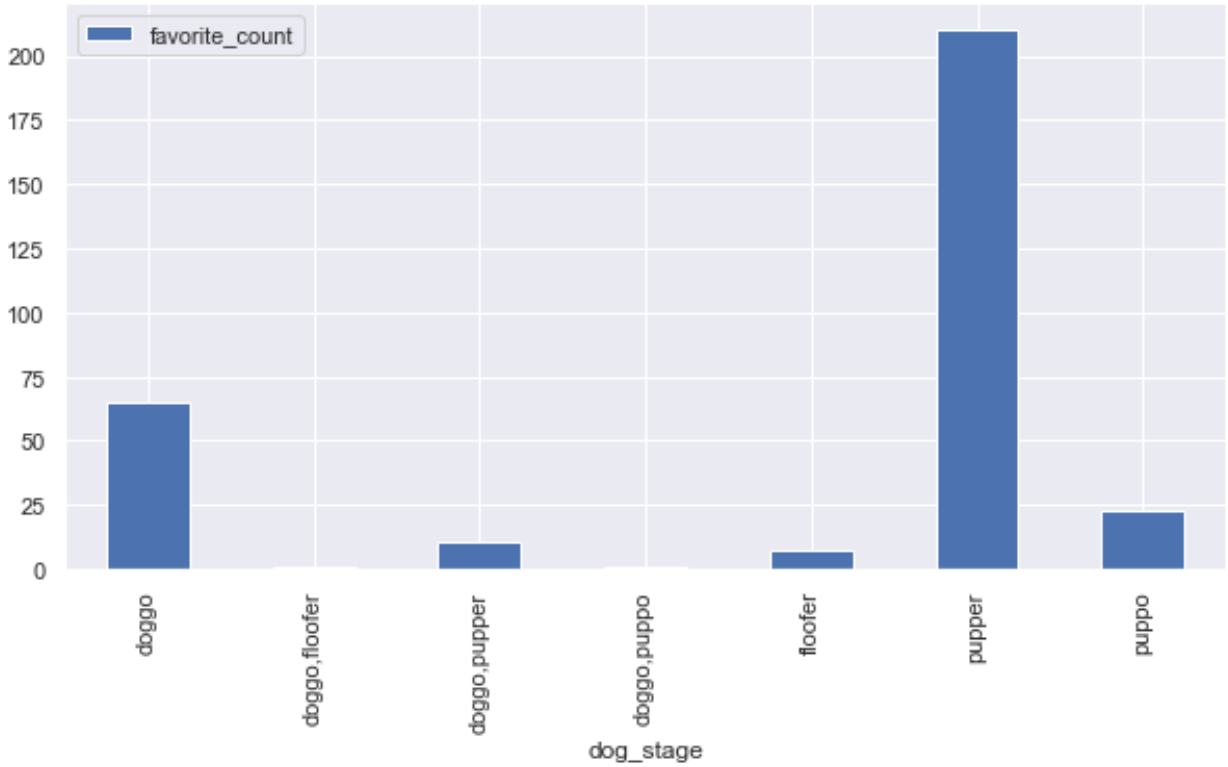
5. What are the dog breeds with the highest tweet ratings?

Answer: Toy poodle breed had the highest retweet counts of all the dog breeds tweeted by WeRate Twitter account with about 47 counts. The visual representation is provided in the bar chart below:



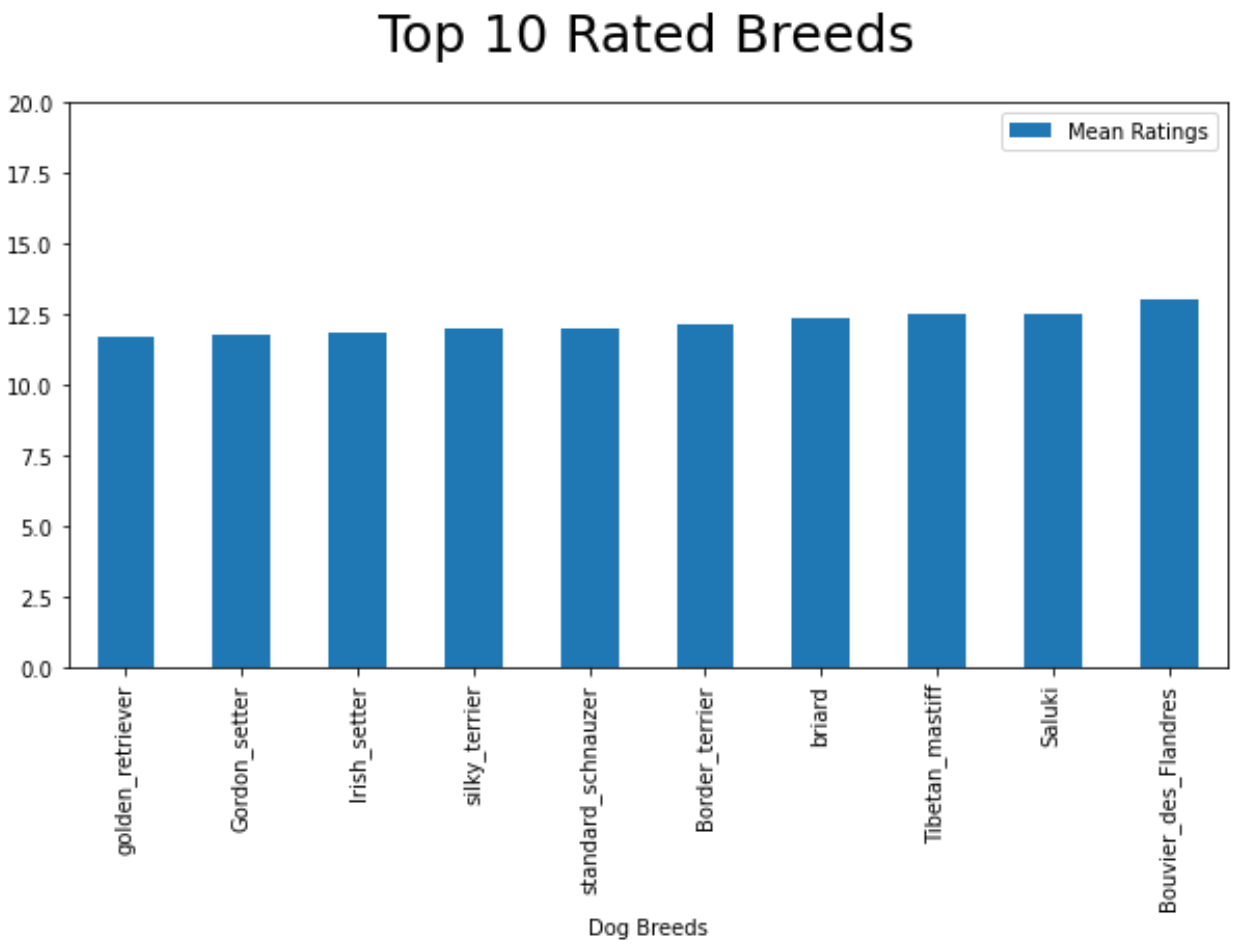
6. **Which dog stage had the most favorite counts?**

Answer: Dogs at the Puppo stage had the highest favorite counts! The visual representation is provided in the bar chart below:



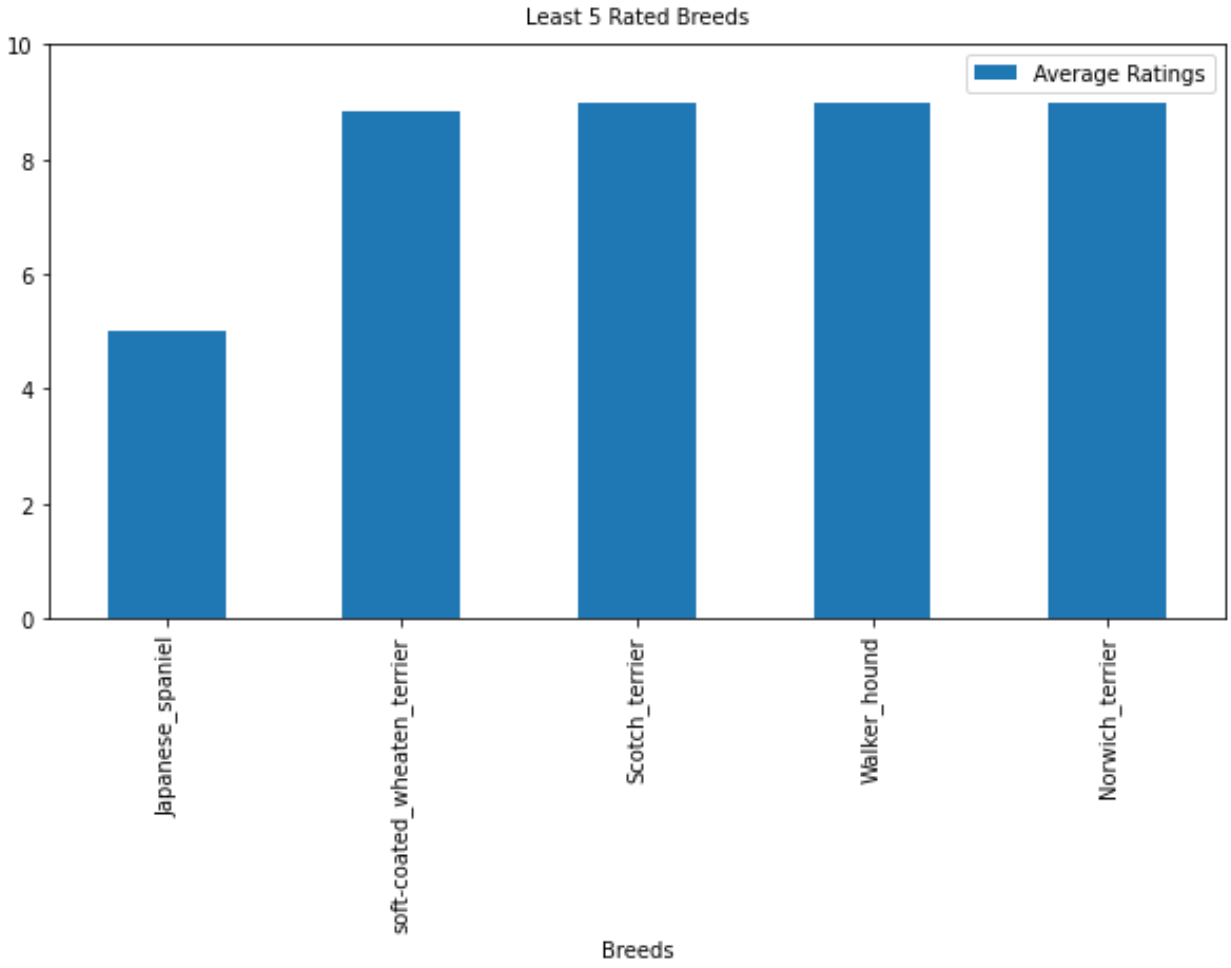
7. What are the dog breeds with the highest tweet ratings?

Answer: Bouvier_des_Flandres and Saluki are the highest rated dog breeds. The visual representation is provided below:



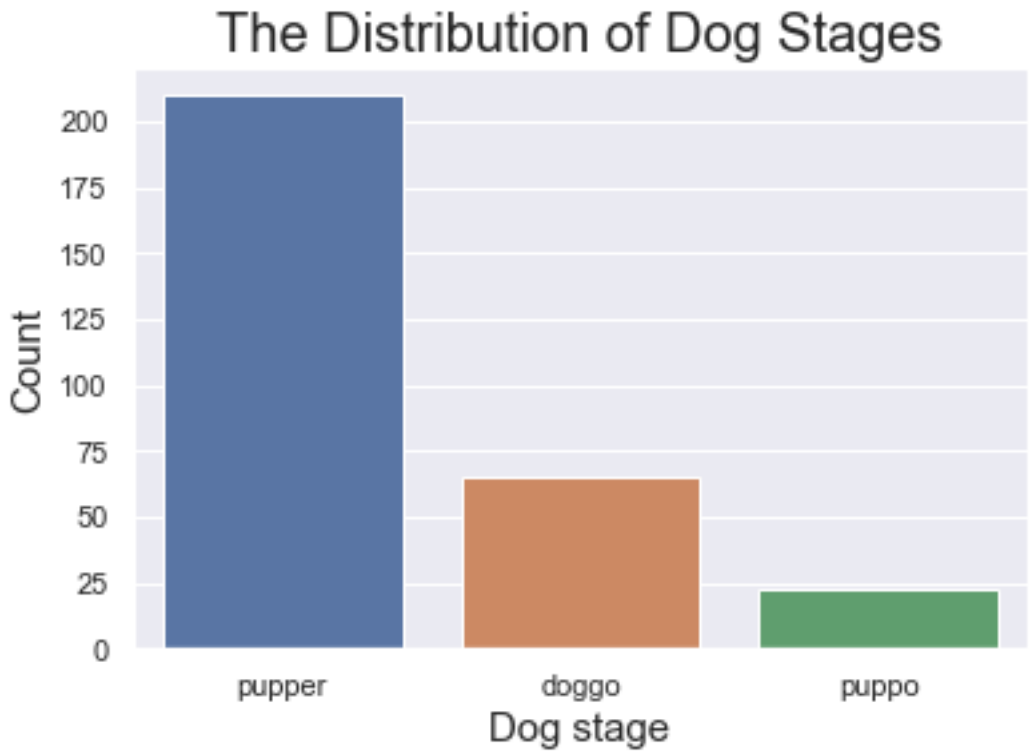
8. **Which dog breeds have the least rated tweets?**

Answer: Japanese spaniel and Soft Coated Wheaten Terrier are the least rated breeds



9. What is the most popular dog breed?

Answer: Pupper is by far the most popular dog stage!



10. What is the mean rating in the dataset?

Answer: A mean rating of 1.17 indicates that the dog tweets from @weRatedogs Twitter account had good ratings.

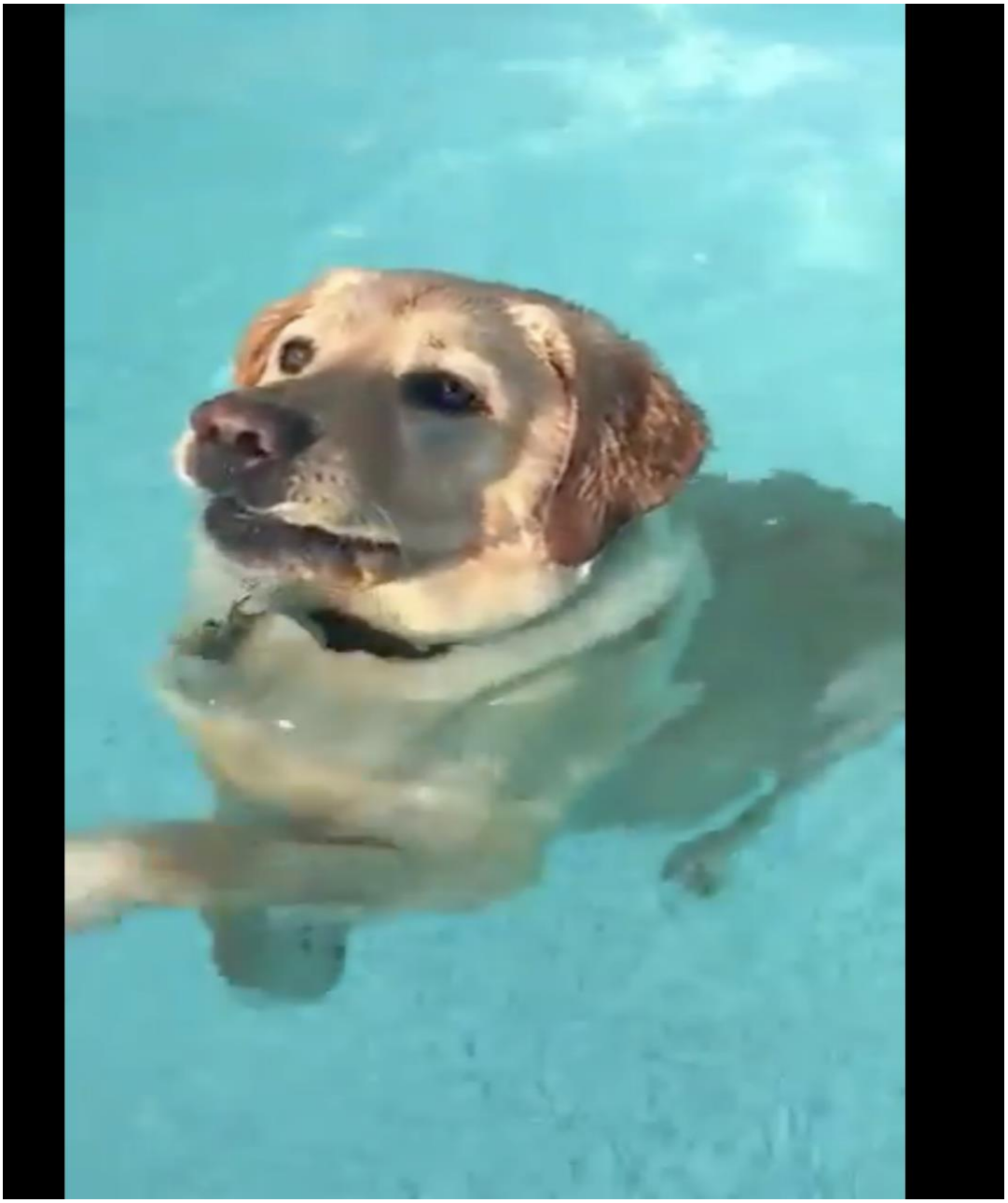
```
count    2057.000000
mean        1.165319
std         4.001623
min         0.000000
25%         1.000000
50%         1.100000
75%         1.200000
max       177.600000
Name: rating, dtype: float64
```

Conclusion:

Without doubt, the WeRate Twitter page analytics is exceptionally good. With an average favorite and retweet counts of 7412 and 2351 respectively, the tweets on the page do not lacking interactions from its over 4 million followers. Its popularity among the followers is also indicated in the rating average of approximately 12/10, which confirms that most of the tweets were highly approved by the followers. Interestingly, the tweet with the id 744234799360020481 was the most retweeted and the most favorited. A look at the tweet hosted at https://t.co/7wE9LTEXC4 shows a picture of a dog swimming in a pool with the following text:

"Here's a doggo realizing you can stand in a pool. 13/10 enlightened af (vid by Tina Conrad".

It is noteworthy to note that the dog in that tweet belongs to the Labrador retriever breed and the doggo stage which were not the most popular in the datasets. So, there must be something very alluring about the dog, perhaps because she was just a good swimmer!



Limitations

The biggest limitation in this project is from the rating values. The max value was 177.6 which was a big outlier from almost 90% of the values. Perhaps the error came during compilation of the values. There was also issue of several missing data in the dog stage column which prevented me for further exploring the distribution of the dog stages.

THANK
YOU