act_report

December 30, 2020

1 wrangling data report

1.0.1 The dataset that we're going to be working on is WeRateDogs Twitter account. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog.

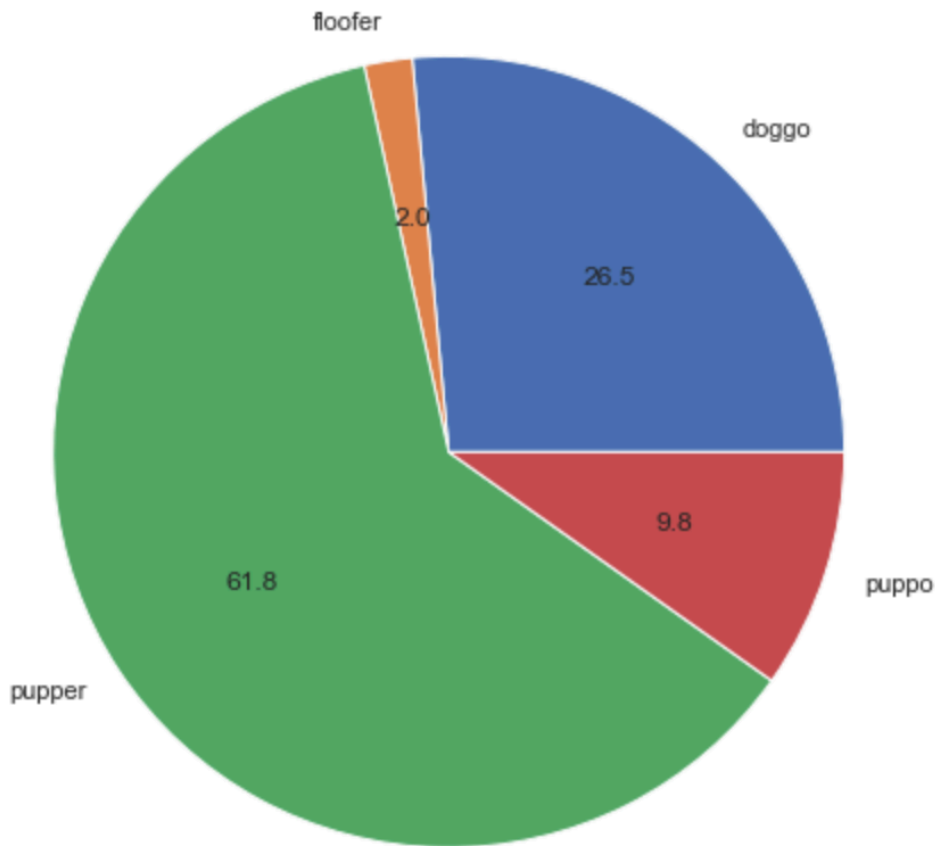### 1.0.2 In this project the three data sources were:

1- twitter-archive-enhanced

2- image-predictions

**3- tweet_json.txt**

1.0.3 wrangling data:

1.0.4 Here are some of the analyzes we came up with during the data wrangling

1.0.5 insight 1:

As we can see in this pie chart, after we calculated the number of existing breeds and counted them, we find that they are 4 dogs breeds, and the most common breed mentioned in the dataset is: pupper with **61%** , then is doggo with **26.5%** , puppo with **9.8%** and the last is floofer with **2%**.
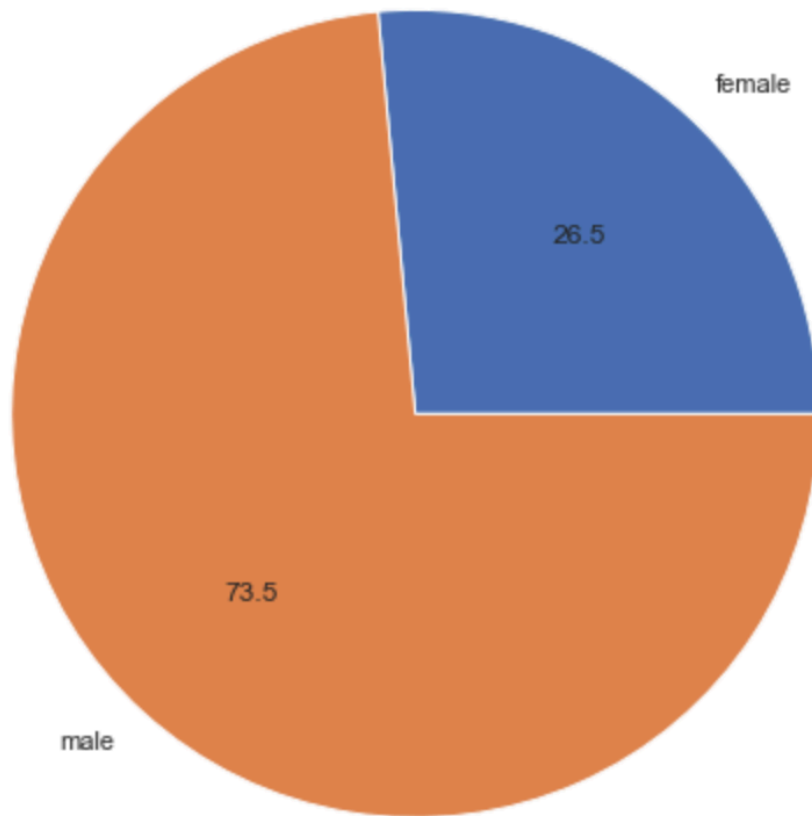
1.0.6 Insight 2:
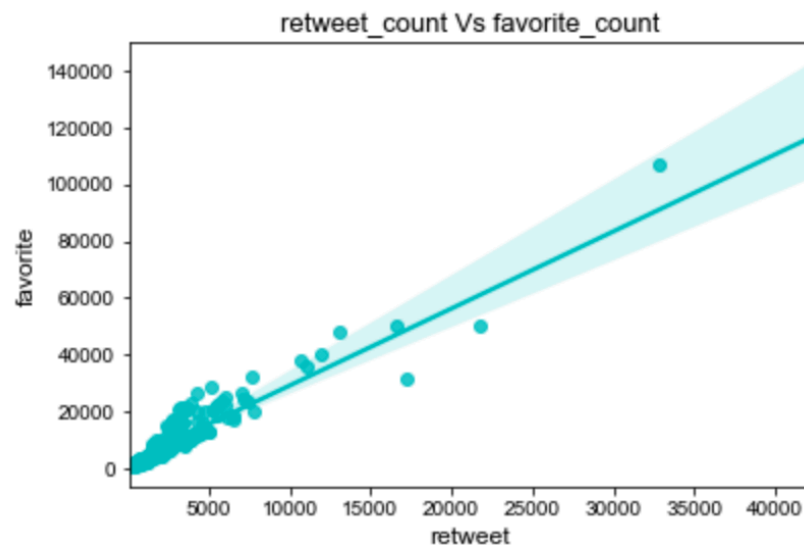
1.0.7 Here in this graph, we observe the relationship of the count of favorites with the count of retweets, although there is a big positive correlation between the retweets and favorites, we see that the density of points is greater when the number of preferences is more than the number of retweets. Therefore, So there are a lot of tweets that have more favorites than Retweets like the following picture:

WeRateDogs™ ✔
@dog_rates

Following

This is Dough. He's a basset hound working undercover as a lab. Can't discuss the mission, but it's very important. 14/10 rare as h*ck

8:28 AM - 11 Mar 2019

9,637 Retweets **112,890** Likes

1.0.8 Insight 3:

In this graph, we see the number of females compared to males, and as it appears that the proportion of males is the highest, and we conclude from this that male dogs are more common and the most willing to adopt.