Analyzing and Visualizing Data

Introduction:

The dataset of the tweet archive of WeRateDogs is used in this project for wrangling, analyzing, and visualizing using Python and its libraries. This account rates people's dogs with funny comments about the dog. We have wrangled this raw data, and it is now ready for analysis and visualization.

Analyzing and Visualizing Data

The requirements to complete this stage are to produce at least three insights and one visualization.

The insights are exploring and visualizing the following question:

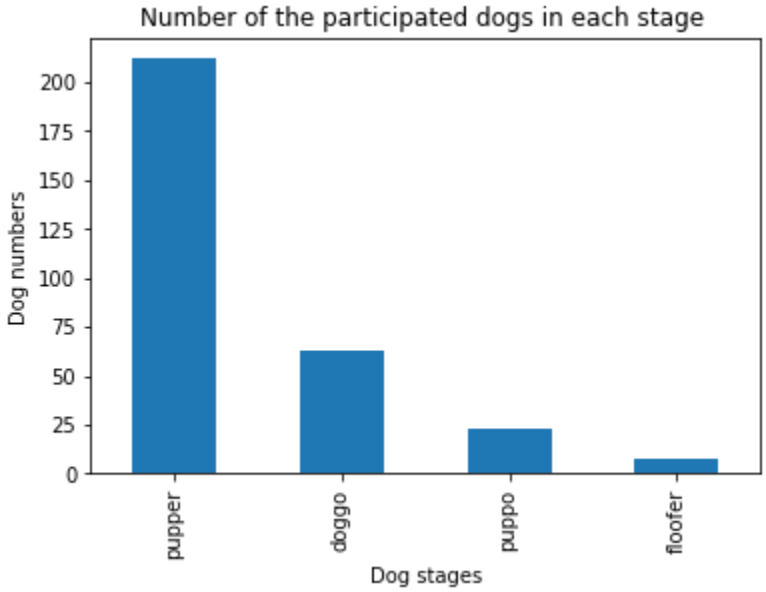
1. What are the most popular dog stages that participated in the tweet by their owners.

2. What are the most popular dog breeds that people like to own?

3. Explain and investigate the relationship between the retweet count and favorite count.

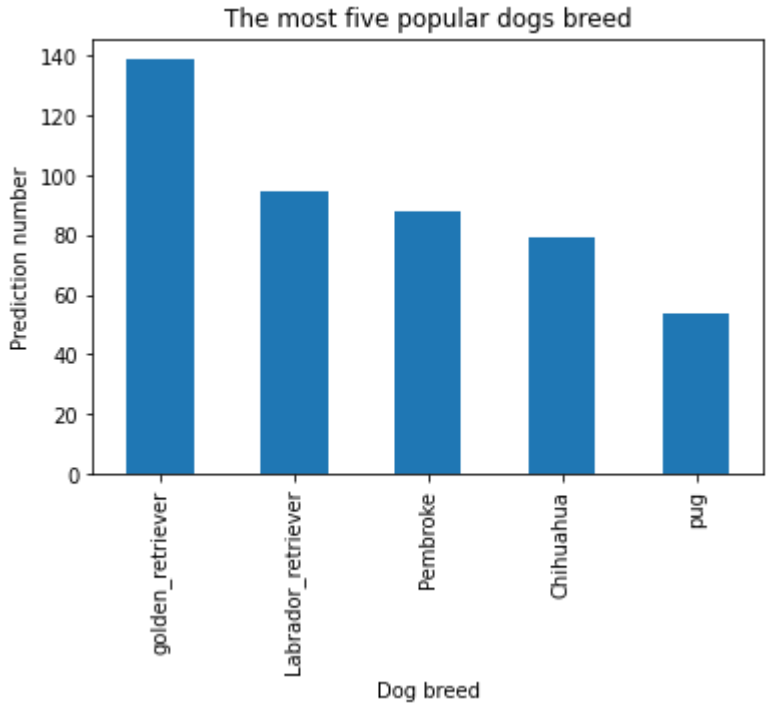
4. Which breeds have the highest and lowest rating?

1. To investigate the most popular dog stages, we used the bar chart to plot the count of each dog stage that participated in the tweet by its owners.



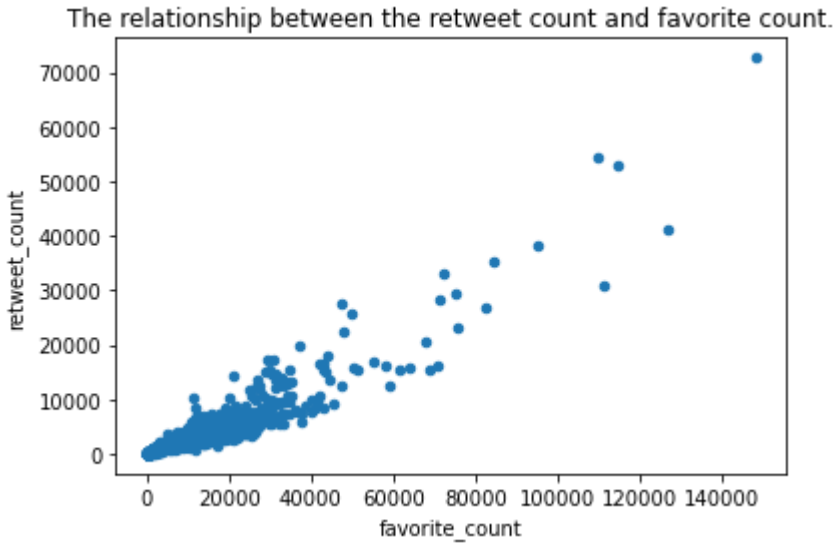
The graph explains that the *pupper* is clearly the most popular of the dog's stages.

2. A bar chart have been used to check the most popular dog breeds that people like to own. The length of breed count is 374, we sliced the first five most popular dog breeds to make it easier to show results and discarded the rest.



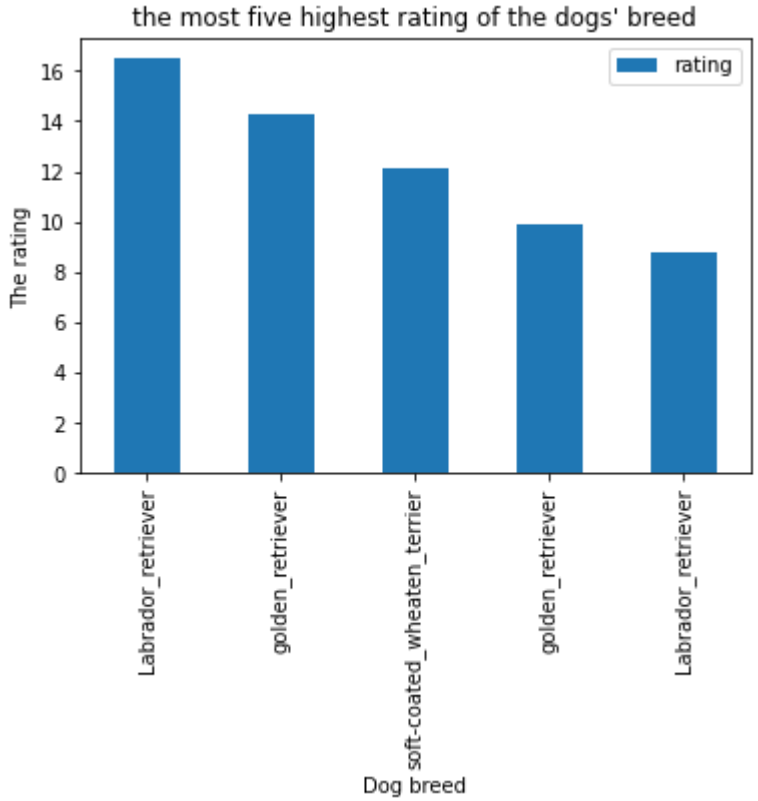
The graph explains that the Golden Retriever is the most popular breed type and the three breeds (Lab., Pembroke, and Chihuahua) have roughly the same rank.

3. To investigate the relationship between the retweet count and favorite count, we plot this relationship using scatter plotting.



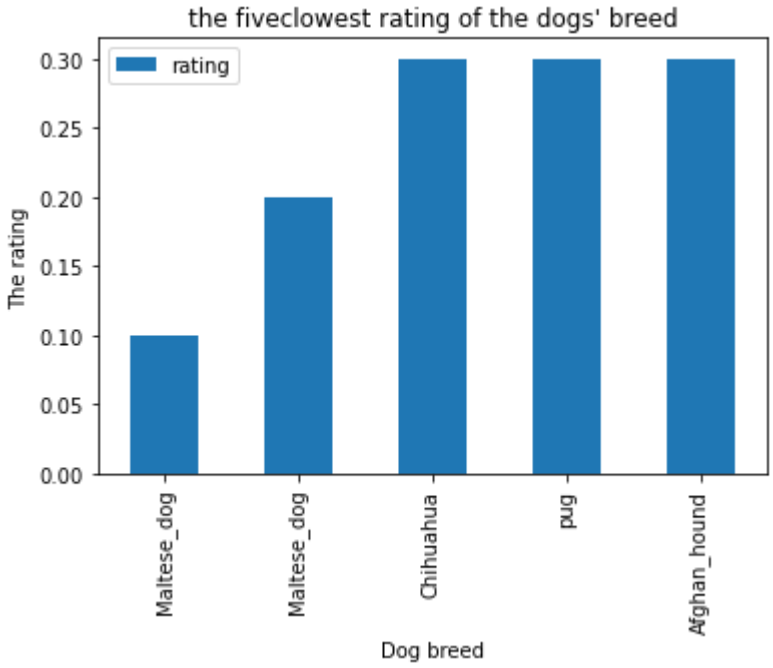
The scatter graph depicts a strong positive correlation between favorite and retweet counts, where the correlation coefficient = 0.9. That is, an increase in the number of favorite leads to an increase in the number of retweets.

4. The following two charts depict which breeds have the highest rating and which have the lowest rating.

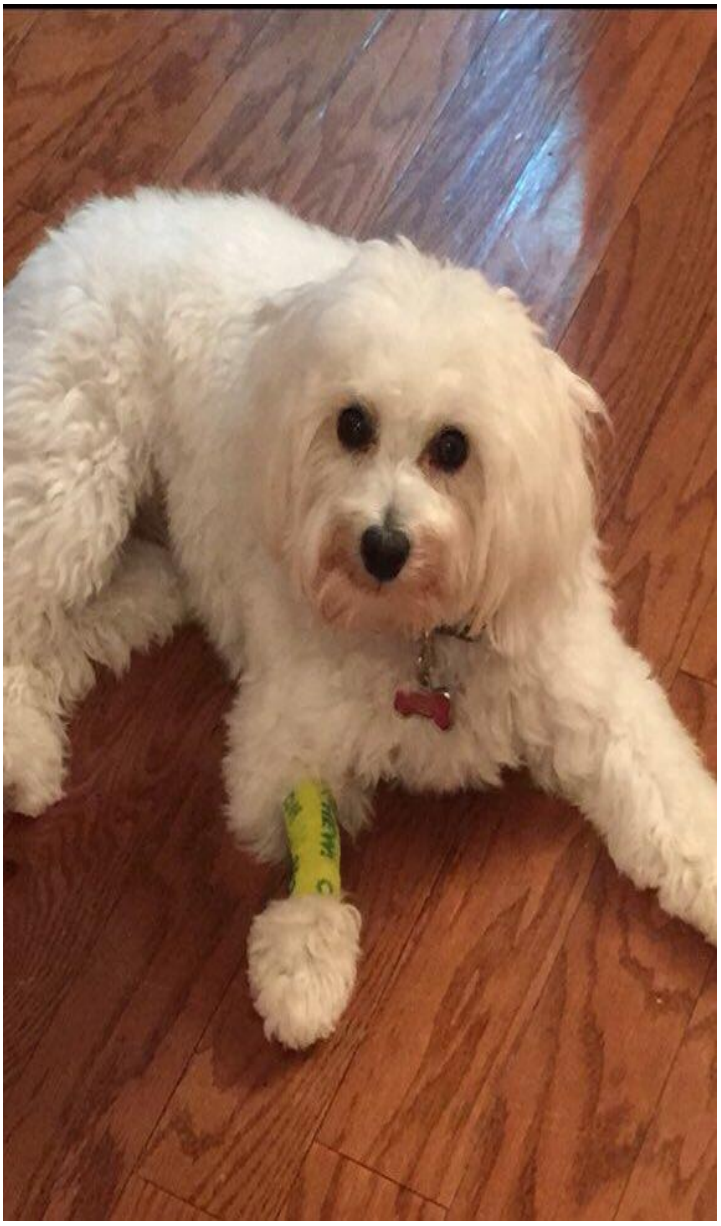


The graph display that the Labrador_retriever has the highest rating and the Gold_retriever comes next in the ranking.

I used the requests library to extract the image and display it.



The graph display the Maltese_dog that has the lowest rating.



Maltese_dog image