

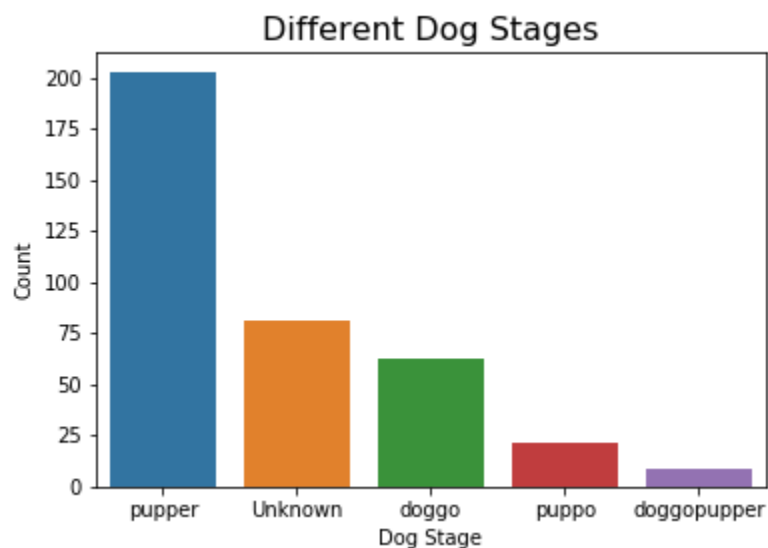
From the analysis made, we asked the following questions which are also answered via visualization using the matplotlib library.

Questions

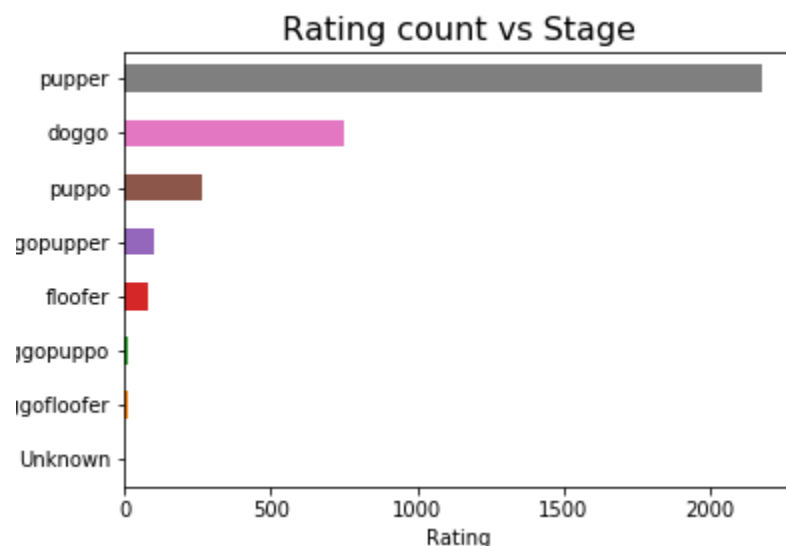
1. Which of the dog stages is highly rated?

From the analysis below, we picked the most common dog stages and plotted it against its count. The most common dog stage found in the data is pupper. It is the most rated dog stage as well.

In order to get the analysis, I first analyzed the first 5 dog stages and found their count, using the `'value_count'`.

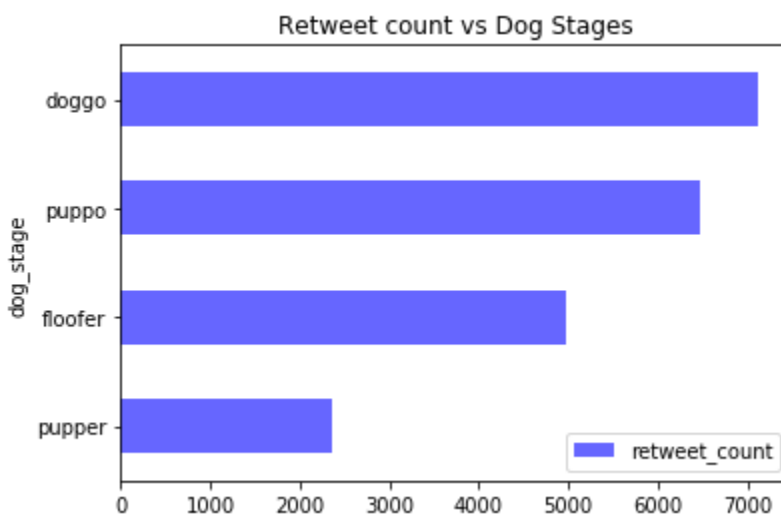
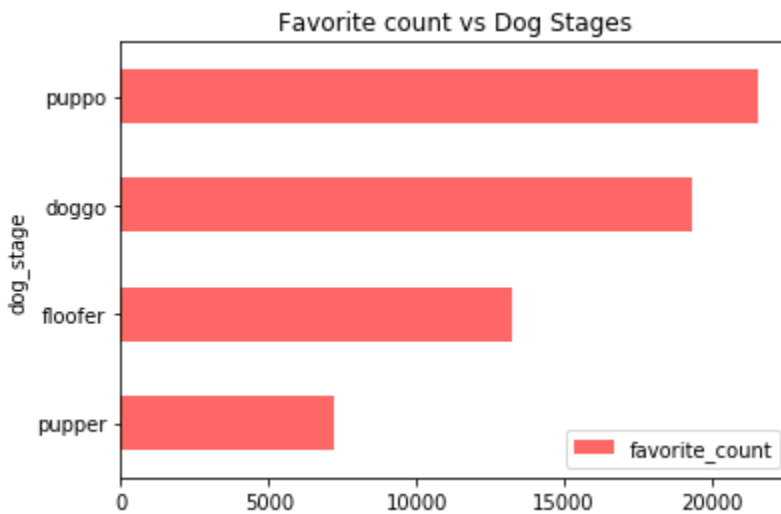


Afterwards, I used the `'group_by'` method to group by `'dog_stage'` and `'rating_numerator'` to analyze the relationship between dog stages and rating.



2. Which dog stage has the most retweet and favorite count?

From the analysis we see that 'Puppo' dog stage has more likes in the tweet data while 'Doggo' dog stage has more retweets. To get this analysis, I queried the data for all the dog stages.



3. Does the retweets and favorite count influence the dog rating?

From the analysis done, yes, retweets and favorites count influence dog ratings.