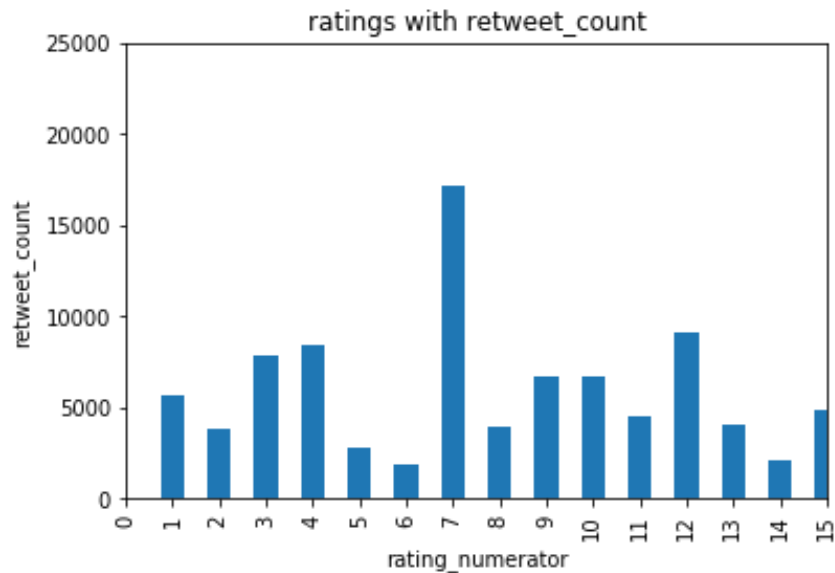
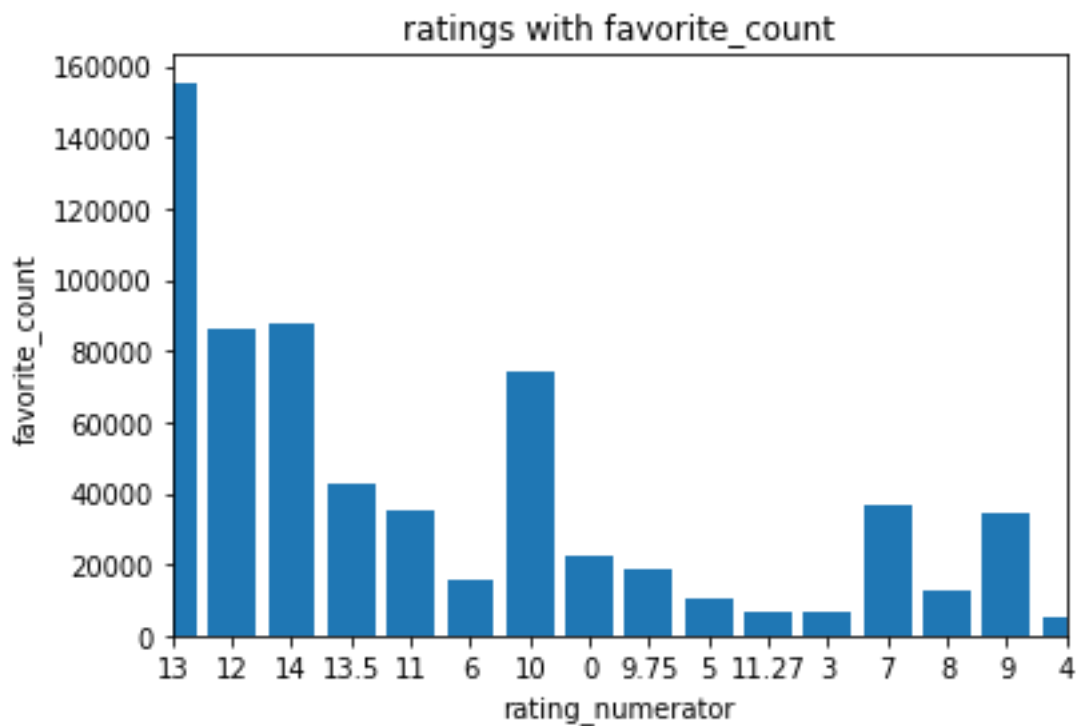


Act report

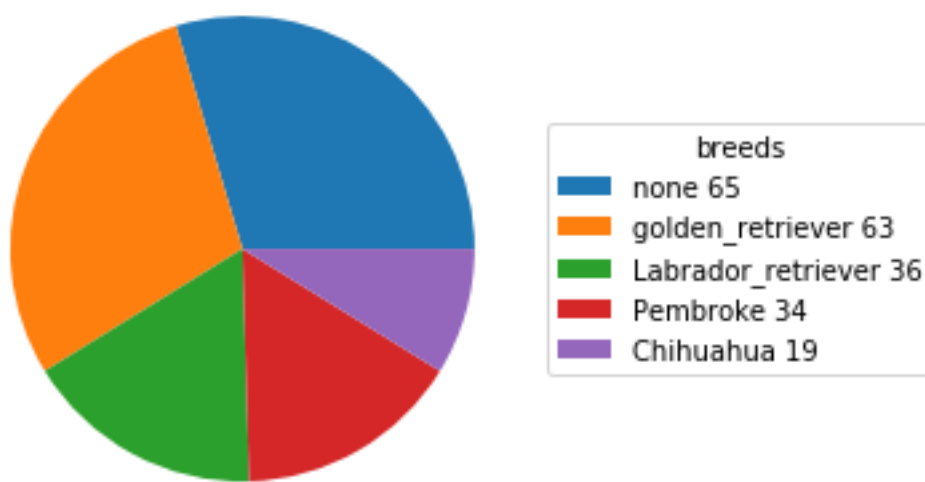
In the first observation you can see that the rating numerator 7 has the highest retweet count.



Interesting is that in the the second observation the rating count 13 had the highest favourite count. an other aspect is that this graph is skewed to the left. Meaning it has shorter bars on the left and longer on the right. Wich is true since this graph has bigger values to the left and smaller values to the right



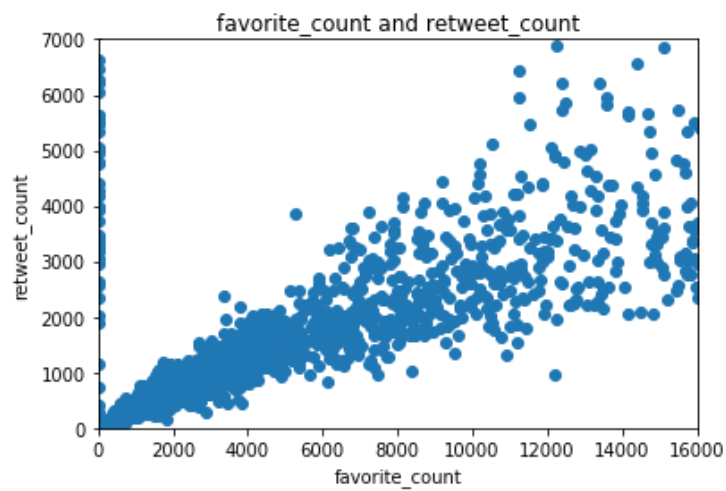
In observation three I took all the rows with a favourite count above the 75% quartile and then looked at the top 5 most common breeds. You can clearly see that there were many missing values, but you can also observe that the golden retriever is the most common breed with a favourite count above the 75% quartile.



If you did not know this is what a golden retriever looks like:

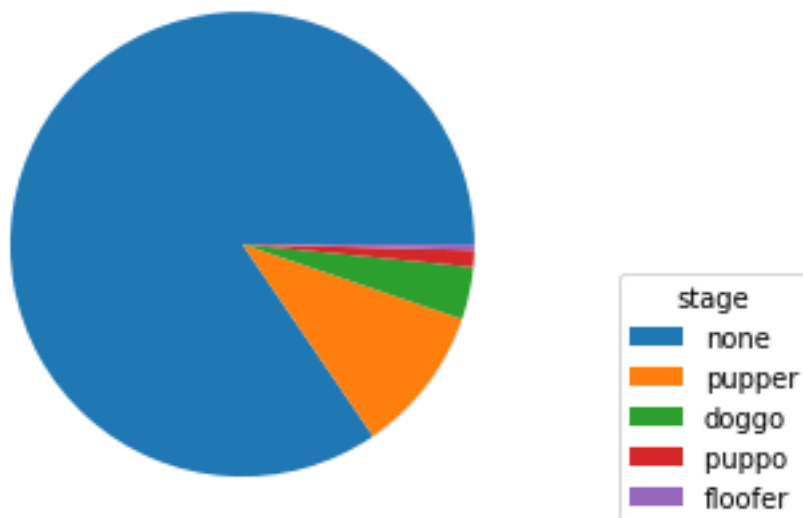


This next observation should tell you if there is a direct correlation between the favourite and like count. You can see that this is a positive scatter plot. Making it clear that there is a direct correlation between the retweet and favourite count.



The next plot/ observation shows a pie plot of the different stages. This value has a lot of none values.

pie plot of different stages

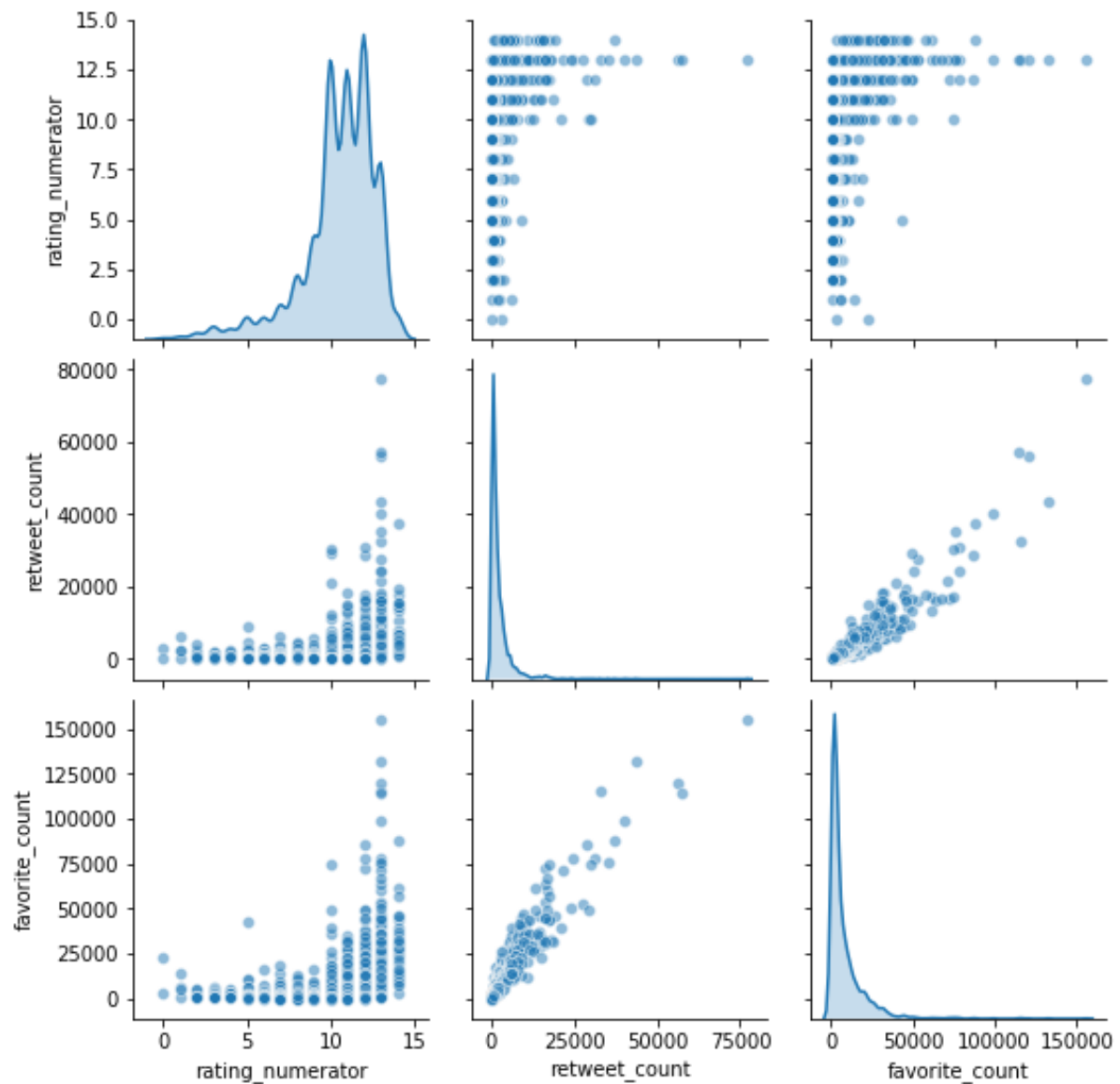


Here are the count of the different stages.

none	1666
pupper	202
doggo	72
puppo	23
floofer	8

In the nine seaborn plots you can see the correlation between 'rating_numerator', 'retweet_count' and 'favorite_count'.

favorite and retweetcount seem to have a correlation as proven before. The plot for rating numerator and favorite count is left skewed. The plots rating numerator and retweet count and rating count and retweet look fairly similar.



In observation 6 I plotted the scatterplot trying to see if there is a correlation between the confidence and favourite count. I also plotted the mean of the confidence and there does not a correlation.

