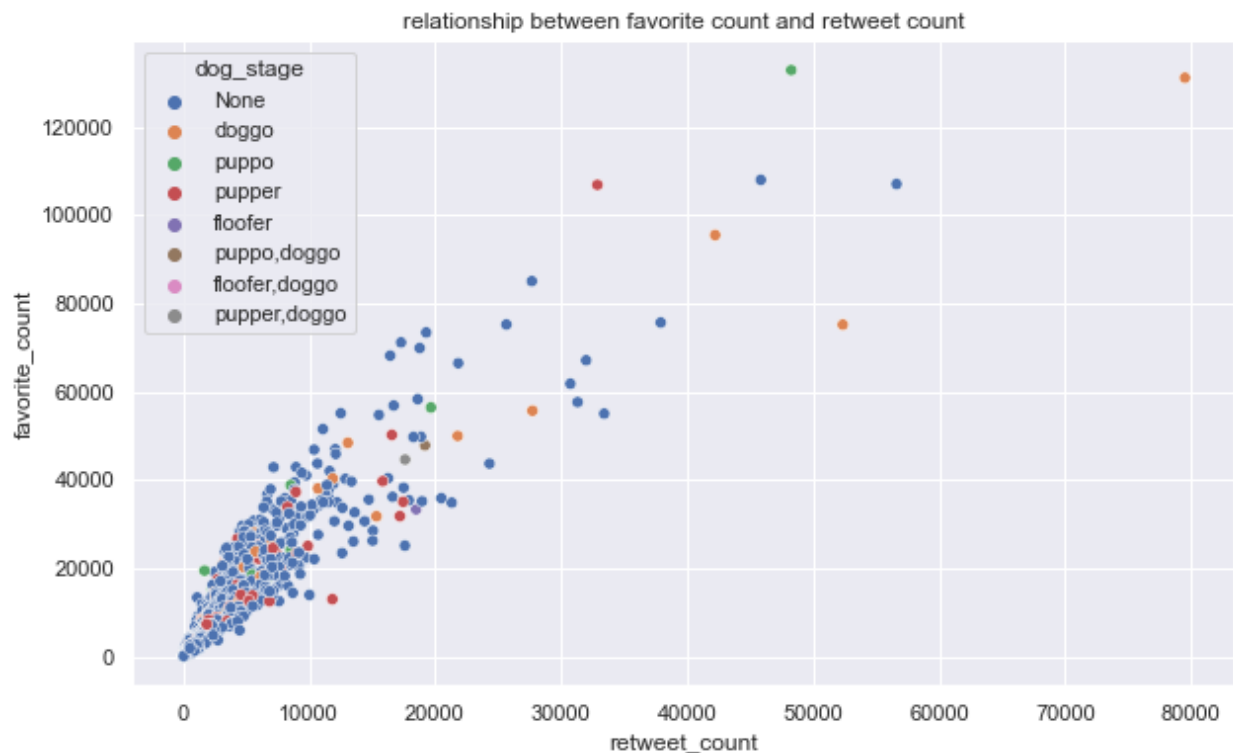


Analysis report

This analysis was done after successful gathering, assessing, cleaning and storing of the different pieces of data required for this project. From the analysis of this data, I was able to get 3 insights. To add on that, there are 3 visualizations that helped to get more clarity on the insights that were drawn from this analysis.

The first insight derived from the analysis is that: There is a positive correlation between favorite and retweet count. This means, when the retweets increase, so does the favorite count.

I arrived at this insight by plotting a scatter plot between the retweet-count column and the favorite_count column. I also used the dog_stage column in the 'hue' parameter of the `sns.scatterplot()` function to show the spread of the different dogstages. The scatterplot is displayed below.



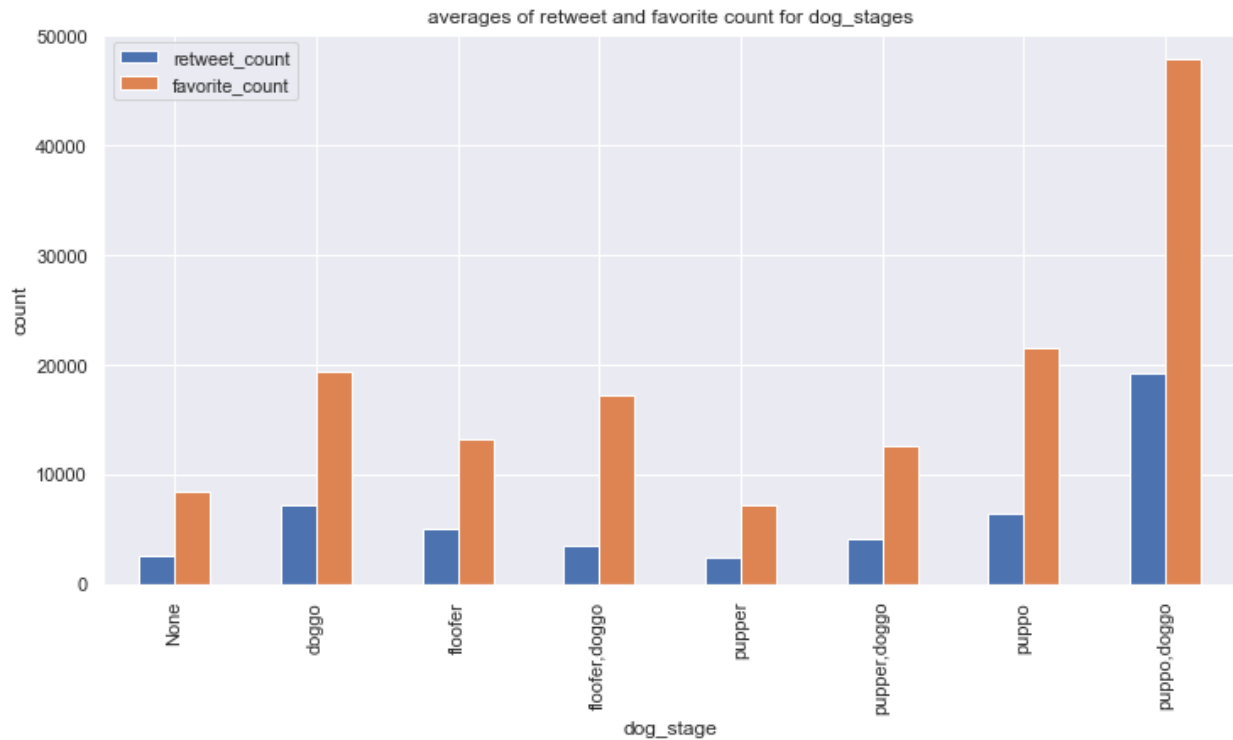
The second insight is that : For tweets whose dog stages have been indicated, it is the pictures that have both puppo's and doggo's together that have the highest favorite and retweet count, this could indicate they are most loved by followers of

the we_rate_dogs account.





To get this, I grouped the data by the dogstages column then got the mean and filtered for the retweet and favorite count. There is a limitation to this insight though, the limitation is that a significant number of tweets did not have an indicated dog stage.



The third insight derived from the analysis is :For tweets whose dog stages have been indicated, it is the pictures with both puppo's and doggos that have the highest rating numerator and denominator. This could mean that the puppo's and doggos are cuter or more lovable.

By grouping the data by dog_stage column and filtering for rating_numerator and rating_denominator and getting the mean, I then was able to plot a grouped bar plot of the dog stages and their rating as shown in the visual below.

