ACT REPORT

I analyzed the 'favorite count column using the pandas '.describe' function and discovered the total count to be 1862 rows, with the mean - 7944.427497. The minimum, 25%, 50%, 75%, and maximum values are 180, 1881.5, 3636.5, 9900.25, and 144054 respectively.

I proceeded to view the image with the least favorite count of 180

IMAGE WITH THE LEAST FAVORITE COUNT

I proceeded to analyze the image with the most favorite count of 144054

IMAGE WITH MOST FAVORITE COUNT

The next thing I did was to repeat the above process for the retweet column, to get the image with the most and least retweets count.

ACT REPORT

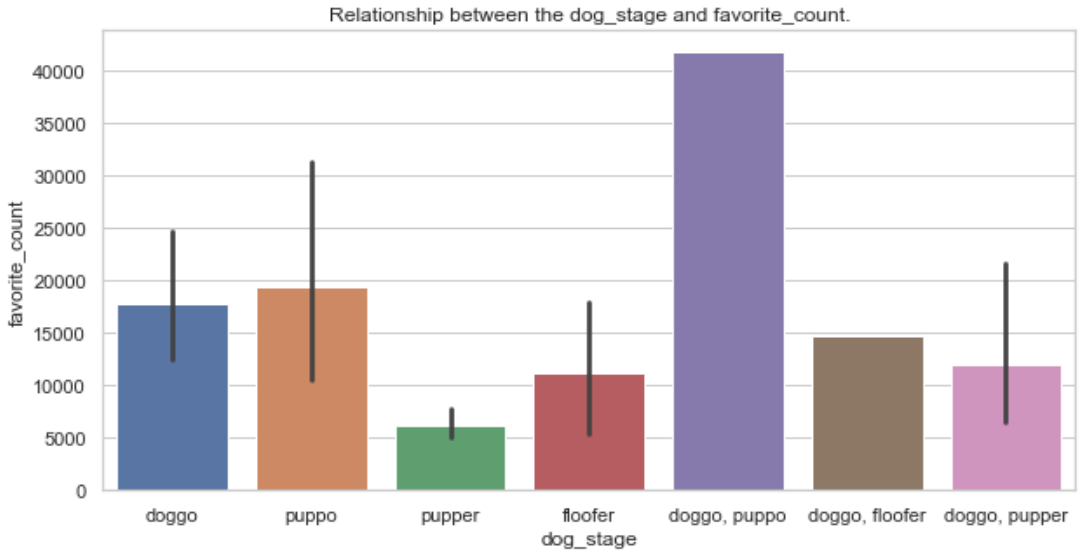
I discovered the total retweet count to be 1862 rows, with the mean - 2291.267991. The minimum, 25%, 50%, 75%, and maximum values are 28, 550.25, 1129, 2614.75, and 70221 respectively.

From my analysis, the image with the most retweets of 70221 was this

While the image with the least retweets of 28 was this

ACT REPORT

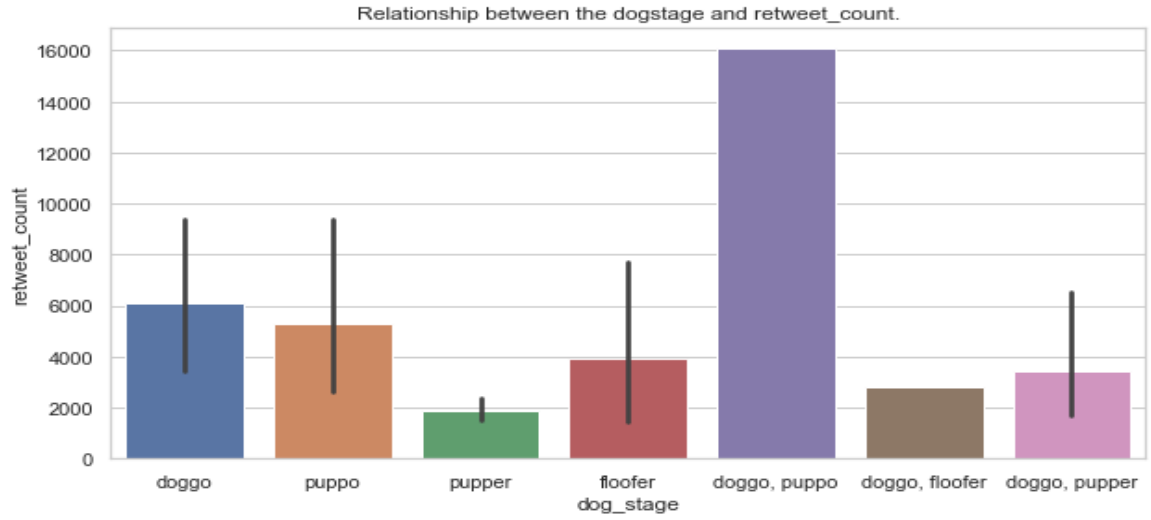
I also analyzed the relationship between the dog stage and favorite count columns. I started by using the pandas '.groupby' function to group the values according to their means, then I plotted a bar chart using seaborn to show the relationship between these two columns.



According to the chart, It was shown that the 'puppo' set of dogs had the highest number of favorite counts while the 'pupper' column had the lowest number of favorite counts.

Visualizing the relationship between the dog stage column and the retweet count.

I also analyzed the relationship between the 'dog stage' column and the 'retweet count' column. I proceeded to use the groupby function to group them according to their means, then I went on to create a bar chart to visualize the relationship between these two columns.



The 'doggo' dogs have the highest number of favorite counts while the 'pupper' dogs have the lowest. The 'doggo, puppo' and 'doggo, floofer' each have only one value that seems to be an outlier and may need to be dropped.