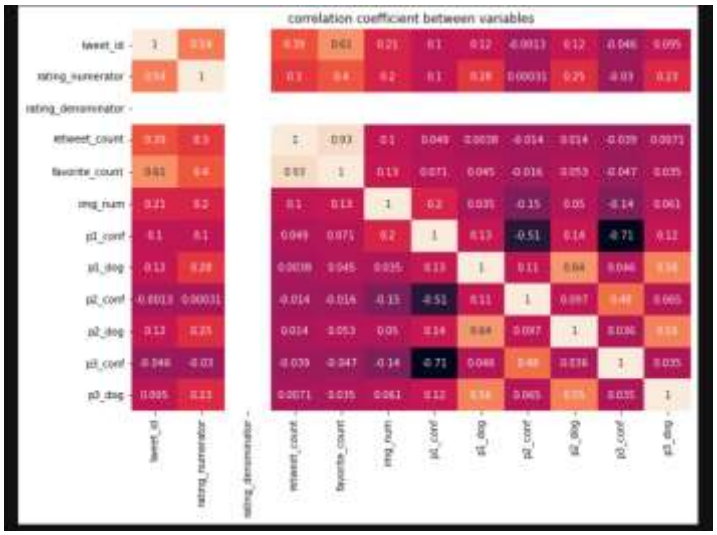
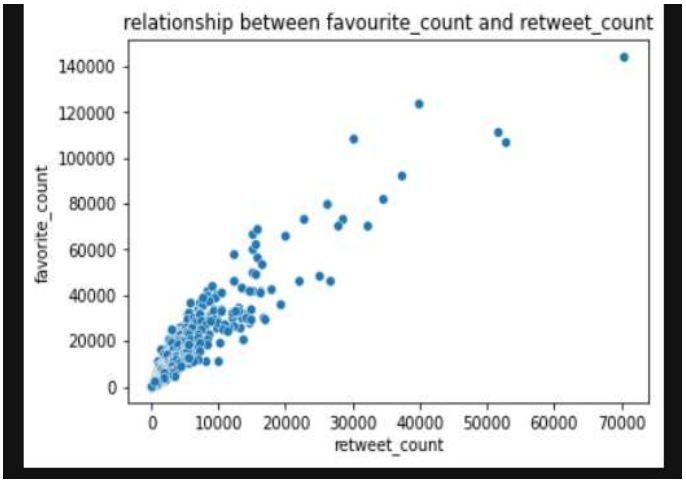
Report on Insight and Visualization

I had insight from the wrangling process which I displayed in the analysis and visualization section.

1. Analyze Relationship between variables

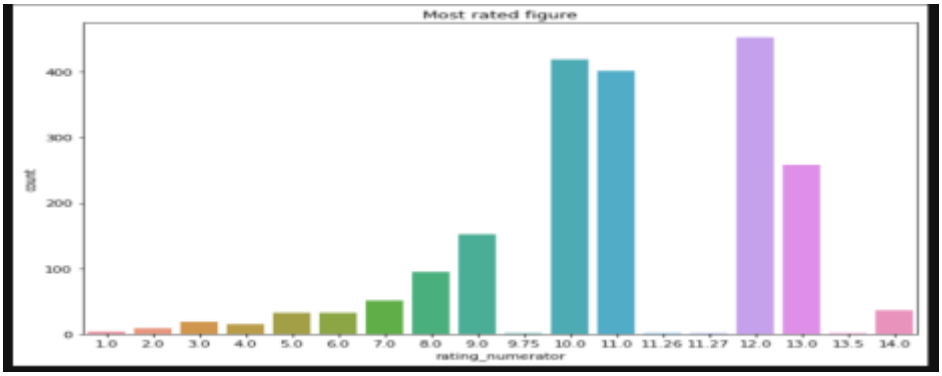




I started off by analyzing the relationship between the variables using the heatmap from seaborn and. corr () method to display the relationship, I discovered the retweet count and favorite count had a significant relationship between which I further displayed with a scatterplot from seaborn. This shows that a good rating comes with increase in both retweet count and favorite count of users.

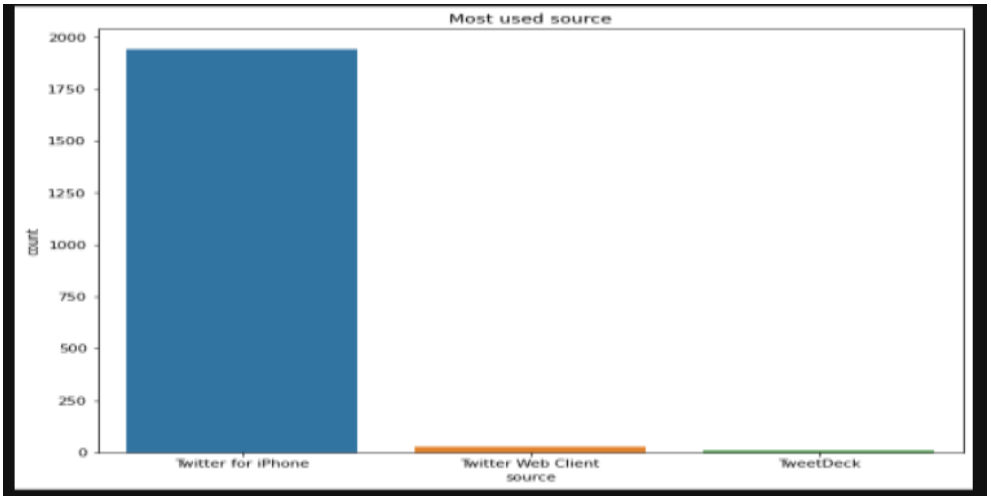
I also analyzed the Dog name column using .value_count() method and i realized the highest name was None.

2. Analyze the most rating figure giving



I analyzed the most rated figure giving to the best dog picture using. countplot () method from seaborn on the rating numerator column and I discovered the most rating figure was 12 / 10 and the average figure was between 10 / 10 and 11 / 10. It was easy for me to conclude after my wrangling process that a very good picture would get a rating between 12-10 even though there were few 13 / 10, 14 / 10.

3. Analyze most used twitter source



Analysis carried out on the twitter source column to ascertain the most used source using. countplot () from seaborn showed 'twitter for iPhone ' was mostly used source from the data gathered.

Further Investigation should be carried out in this column to know if there are limitations in the user interface of the twitter for android app and twitter web for using the retweet and favorite (like) features to reduced biased conclusion on the source.