# Act Report

Rating and Number of Favorite

I was first interested in the relationship between 'rating' and 'number of favorite'. I created a boxplot and found out that the average number of favorite increases as the rating increases to certain extent(numerator rating of 14) I could say that people tend to give favorite to those tweets with higher score.

Boxplot between rating_numerator and favorite_count
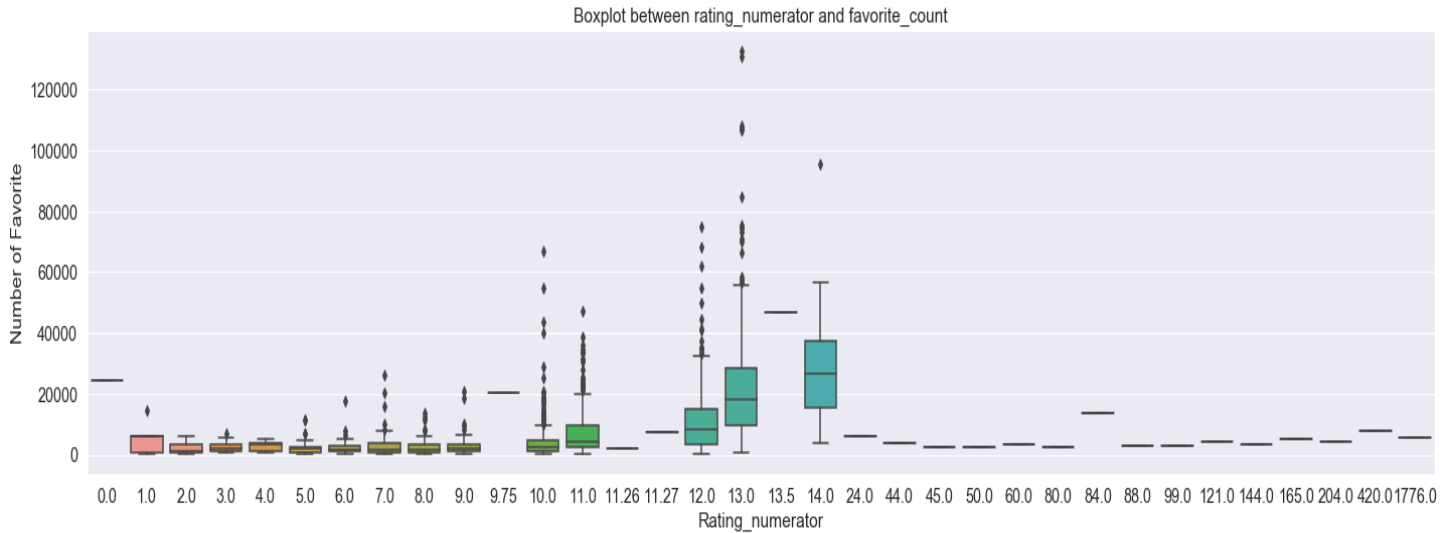
Rating and Number of Retweet

So I was also interested in whether the number of retweet has any relationship with the rating. And I could get the similar result from their boxplot as well. It turned out that highly rated tweets were more likely to be retweeted from the people


Boxplot between rating_numerator and retweet_count
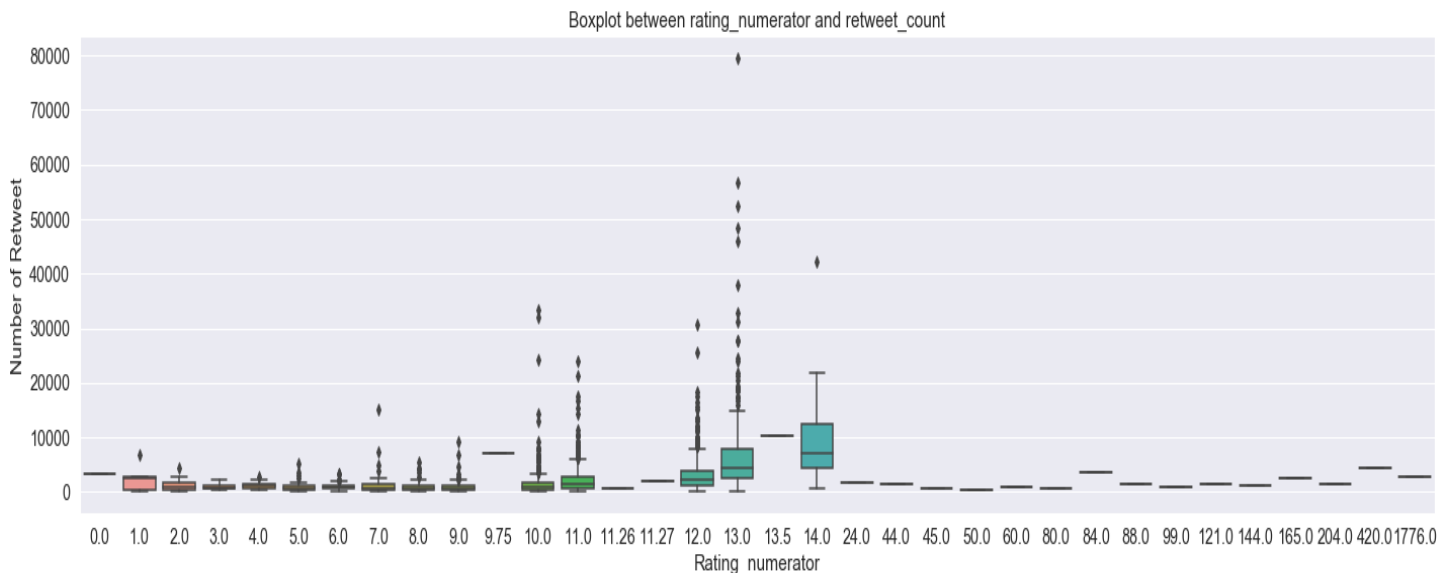
Scatterplot between 'favorite_count'' and 'retweet_count'

So I looked for the correlation factor between 'favorite_count' and 'retweet_count' with 'np.corrcoef' function and created the scatterplot. I could get the correlation factor of 0.91 which was very high and the scatterplot showing the positive correlation relationship.
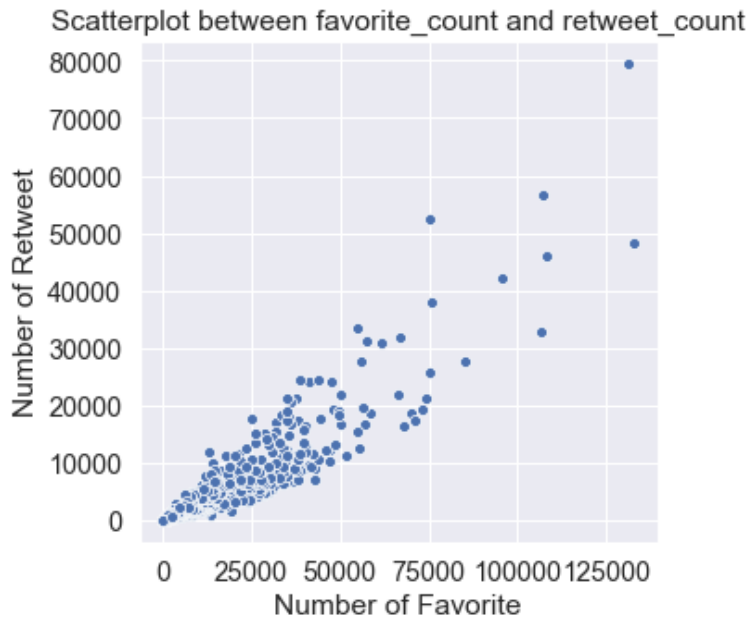

Lineplot between 'stage'' and 'favorite_count'

I was interested in whether there is any preference in dog stages. So I created line plot between 'stage' and 'favorite_count'. There was a notable increase in the number of favorite_count in the tweet with 'doggo/puppo' stage. The other stage like 'doggo, doggo/floofer, doggo/pupper, floofer, pupper, puppo' all had similar number of favorites. But only the stage 'doggo'puppo' had almost doubled or tripled number of favorites. It was a very intriguing result that people like the dogs that both have traits of **doggo and puppo**.
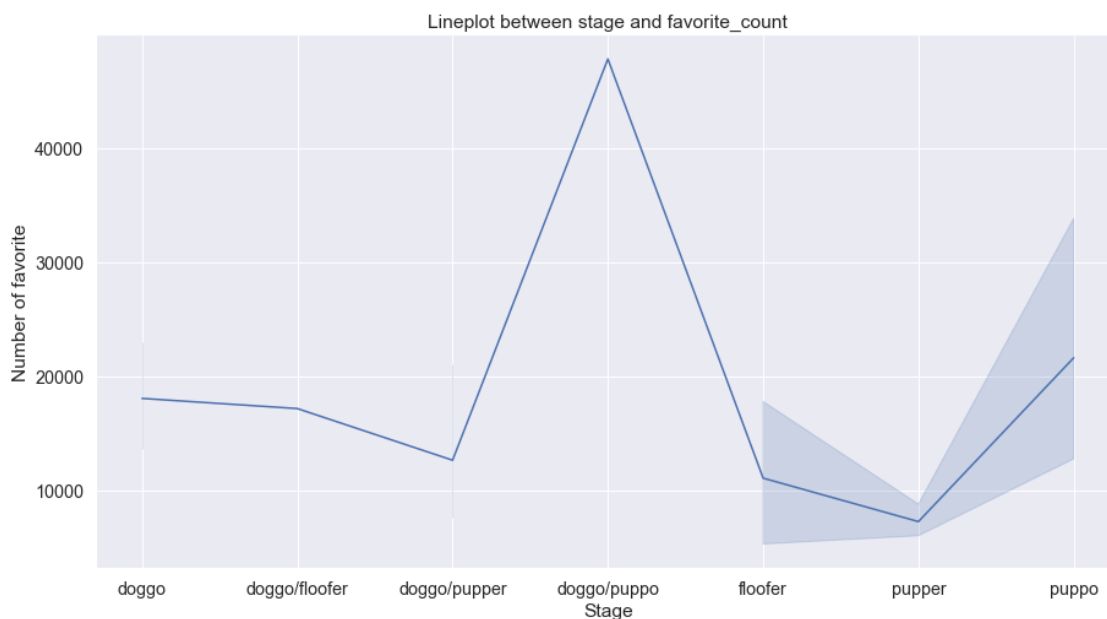
Top 10 Names

I looked for the top 10 popular dog names simply by using .value_counts() function.
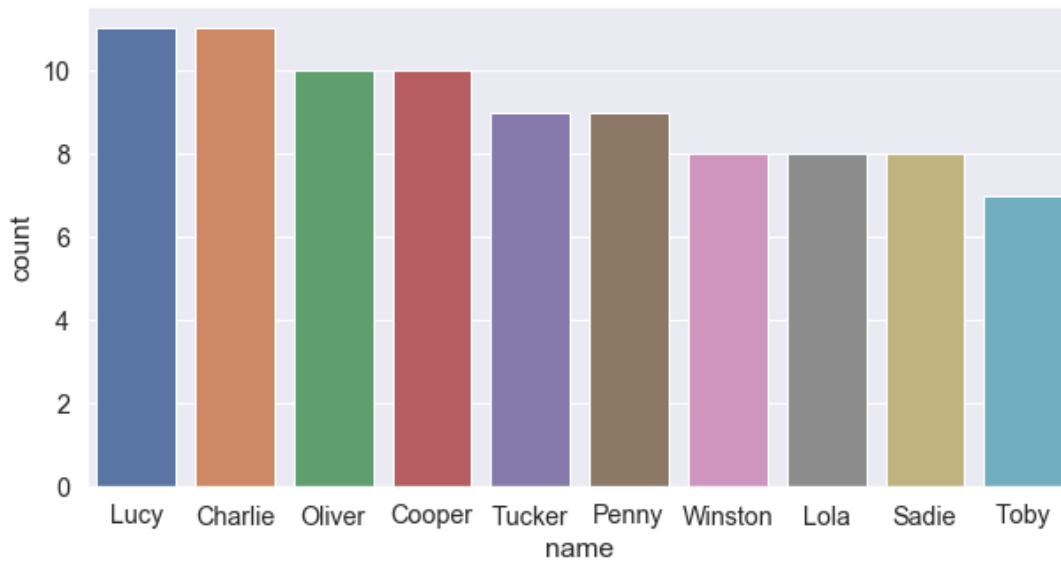
It can be found that the Top 10 popular names are as follows: **['Lucy', 'Charlie', 'Cooper', 'Oliver', 'Tucker', 'Penny', 'Winston','Sadie', 'Lola', 'Toby'].** This result was very unexpected since I have never seen such a dog named 'Lucy' every in my life (Yes I have a very narrow relationship with people)

Countplot of Timeline(hour)

Suddenly I was interested in the activity hour of WeRateDog Account owner. So I created the countplot of his tweet hour. Interestingly, He was most active during **(00:00 ~ 03:00),** so mostly after midnight. Then his activity was very low during (07:00 ~ 14:00). We can infer that he may be sleeping or going to work during that hour.
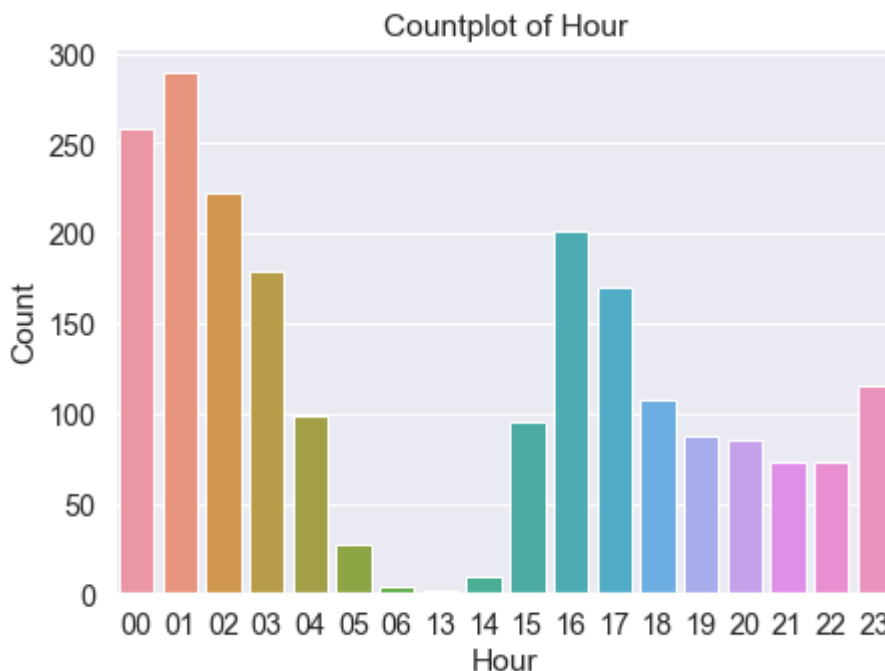
Since the Account owner still post the tweet until 05:00 am in the morning, there is a high possibility that he is a 'night owl' guy.