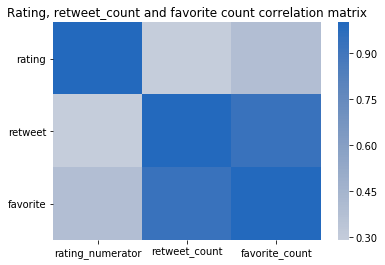
# WeRateDogs Twitter Data Analysis Findings Summary

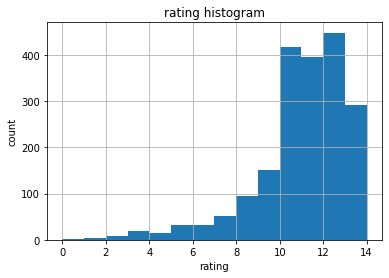
## Rating\_numerator, retweet\_count and favorite\_count Distribution and Relationships

First, take a high-level view of the owner’s rating ( rating\_numerator) , retweet\_count and favorite count:

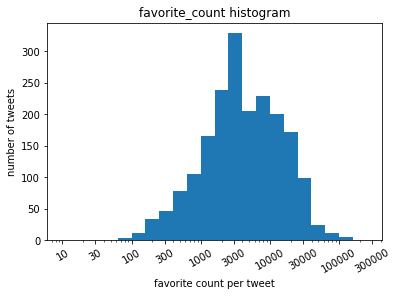
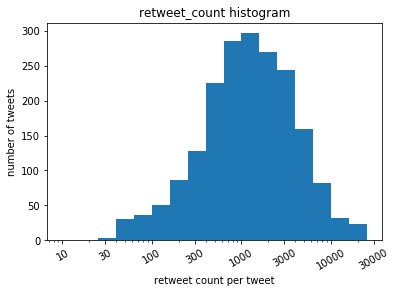


Correlation matrix heatmap shows, that owner's rating is moderately correlated with both retweet\_count and favorite\_count. Retweet\_count and favorite\_count are highly correlated with each other. All the correlations are positive.

For each of the above metrics, a histogram is drawn to depict the metric’s distribution.

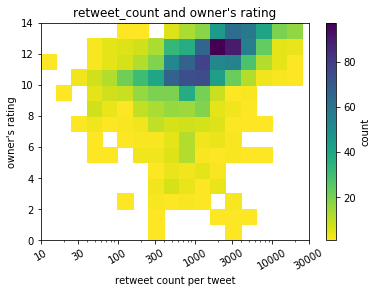


The owner's ratings follow a left-skewed distribution, while majority ranges from 10 to 14.

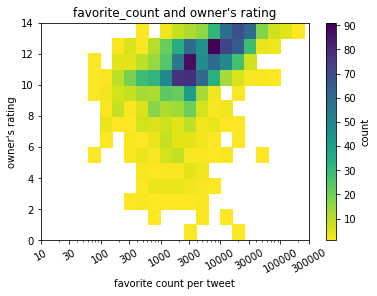


Retweet\_count and favorite\_count both follow log-normal-like distributions. Retweet\_count distribution peaks between 1000~2000, while favorite\_count distribution peaks at around 3000.

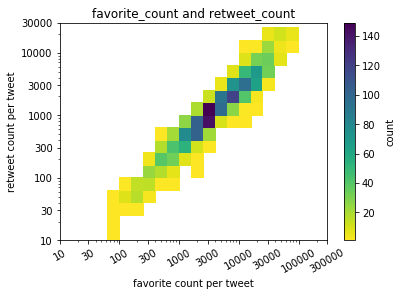
Now look into the 2d histogram of each pair out of the 3 metrics:



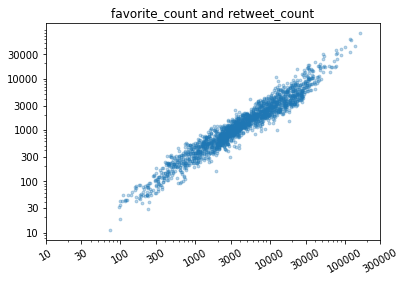
Owner's rating is moderately positively correlated with retweet\_count.



Owner's rating is moderately correlated with favorite\_count.



Add another view of favorite\_count and retweet\_count by scatter plot:



Retweet\_count is strongly correlated with favorite\_count.

## Dog Breeds and Popularity

The most tweeted 10 dog breeds are (starting from highest):

'golden retriever',

'Labrador retriever',

'Pembroke',

'Chihuahua',

'pug',

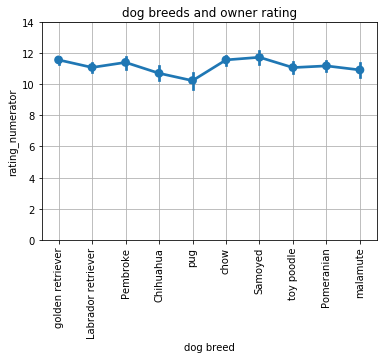
'chow',

'Samoyed',

'toy poodle',

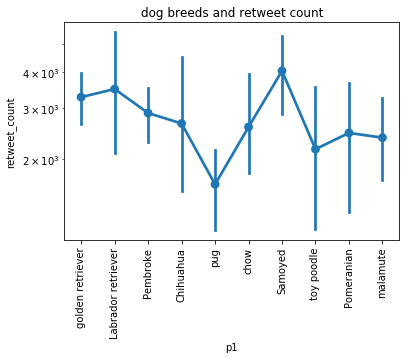
'Pomeranian',

'malamute'



Owners in general give high ratings to their dogs, often above 10 (out of 10). All of the 10 most tweeted dog breeds get more than 10 out of 10 ratings from their owners, averaged over each breed. Out of them, Samoyed gets the highest average rating 11.74, followed closely by Golden Retriever with average rating of 11.57. The lowest rated among the top 10, the Pug, gets 10.24 in average.

Readers’ responses towards dog breeds can be represented by retweet\_count. Because of the log-normal distribution of retweet\_count as mentioned before, the y-axis of the following chart is on log scale.



The most retweeted dog breed among the top 10 popular breeds is Samoyed, while the least retweeted breed is Pug. This is consistent with the owner’s rating towards these breeds.