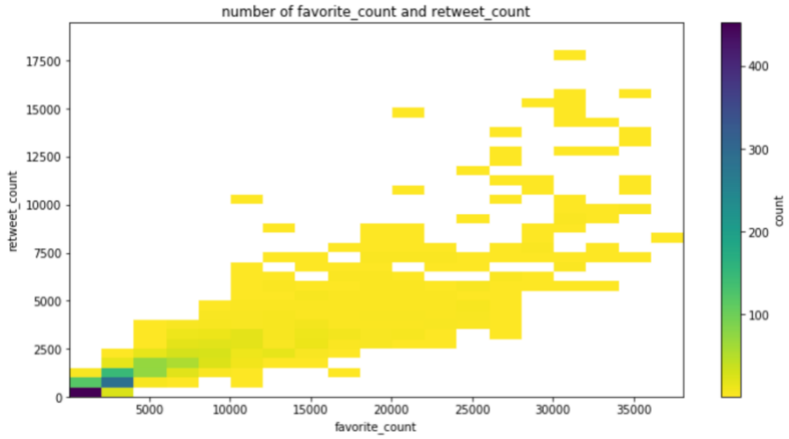
Act Report

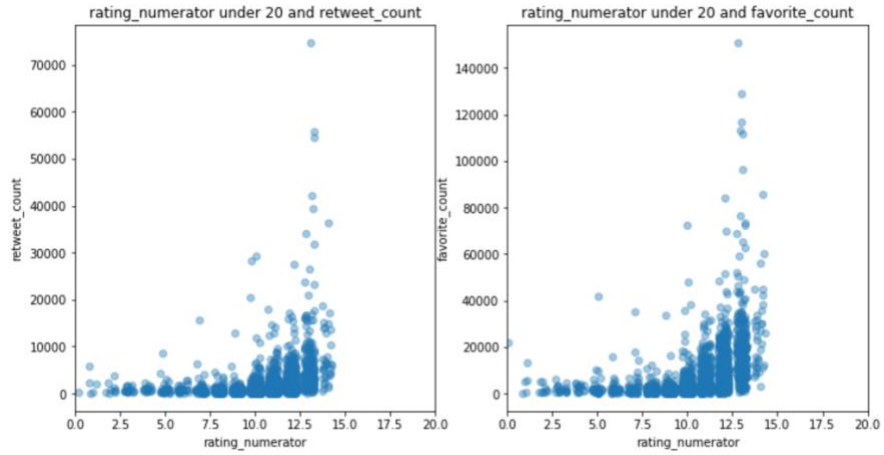
Relationship between favorite_count and retweet_count

From the histgram, we can see the favorite count and retweet count have a positive relationship. When looked at the stats of these two factors, we can see the mean of favorite count is around 3 times of retweet count. Although the max value for these two factor are more than 150,000 and 70,000, 75% of the favorite count is less than 10,000 and 75% of the retweet count is less than 3,000. The boxplot can tell how spread these two factors are, as there are many outliers. Same story as the heatmap, the color is all light after (10000,2500).



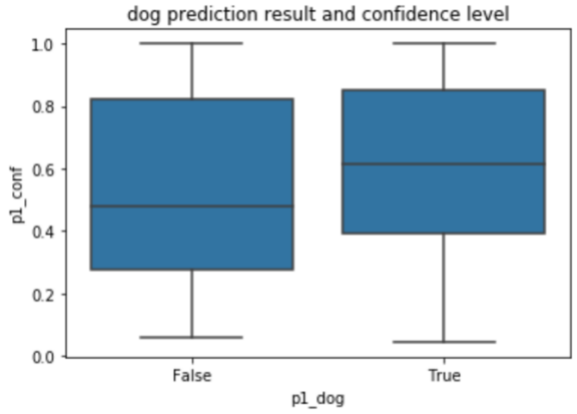
How rating_numerator affect favorite_count and retweet_count?

One of the features that make Dog_rating popular is their rating score. It caught my eye when the rating_numerator is way high. Although denominators are normal, 75%+ is 10, the numerator has a max value of 1776. Almost all of the numerators are higher than 10. Does increasing numerator help the tweet to win more favorite or retweet? I made several scatter plots and they are showing that favorite or retweet count surges when the numerator goes up. But when the numerator is bigger than 15, the tweets don't get higher fav/retweet count any more.



Relationship between p1_dog and p1_conf

When I made a boxplot for prediction of dogs' images, the 'True' result has a higher confidence level than the 'False' result. And if compared the mean, p1_conf > p2_conf > p3_conf. p1 looks good and doesn't have many outliers as p2 and p3.



p1_dog, year and favorite_count

The mean of favorite_count increased significantly during the past three years, from 2015 to 2017. There's no big difference between false and true predition. But 'false' prediction seems to have larger deviation around the mean, which means the 'false' predition (the pictures are usually not focusing on the dog, but other stuff, like an orange) some people don't like this kind of pictures while some people enjoy them than other normal pictures with dog-focused.