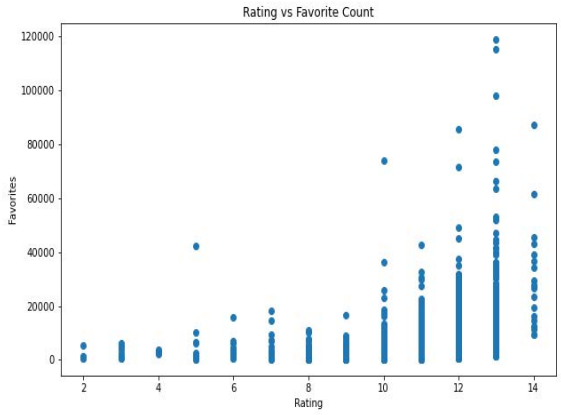
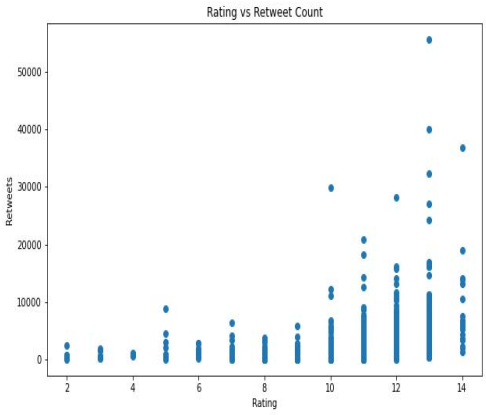
In the analysis phase, some interesting questions to look at are whether or not there is a relationship between the dogs with the highest favorite count and the dogs with the highest rating. It can be clearly seen by visualizing a scatter plot that as rating increases, so does popularity and favorite count. Similarly, as the rating increases the number of times a tweet gets retweet increases. However, at a rating of 13 there are some observations with higher favorite and retweet counts than the highest favorite and retweet count of a rating of 14, so it looks like popularity peaks at the rating of 13.

To figure out which dog was the most favorite, the data with the dog image predictions was used. Dogs with a probability of greater than 0.95 that their image correctly identifies their type were used. This probability is then set lower to 0.90 to increase the number of dogs looked at, and is refitted into the application to allow for higher error rate. The first most favorite dog type is a samoyed, with probability of 0.95 of being a samoyed and an image of the dog is shown in analysis. The second favorite dog type is a french bulldog, with probability of 0.90 of being a french bulldog and is shown in the analysis.

Finally, the most popular dog names on WeRateDogs are analyzed, and this includes 'Cooper', 'Lucy', 'Charlie', and 'Oliver'! These four dog names appear 10 times in our dataset. Next is ‘Tucker’ and ‘Pennie’ (9 dogs), followed by ‘Sadie’, ‘Winston’, and ‘Lola’ (8 dogs).



Rating vs Favorite Scatterplot



Rating vs Retweet Scatterplot