Summary

1. How good the image prediction algorithm was to predict dog images?

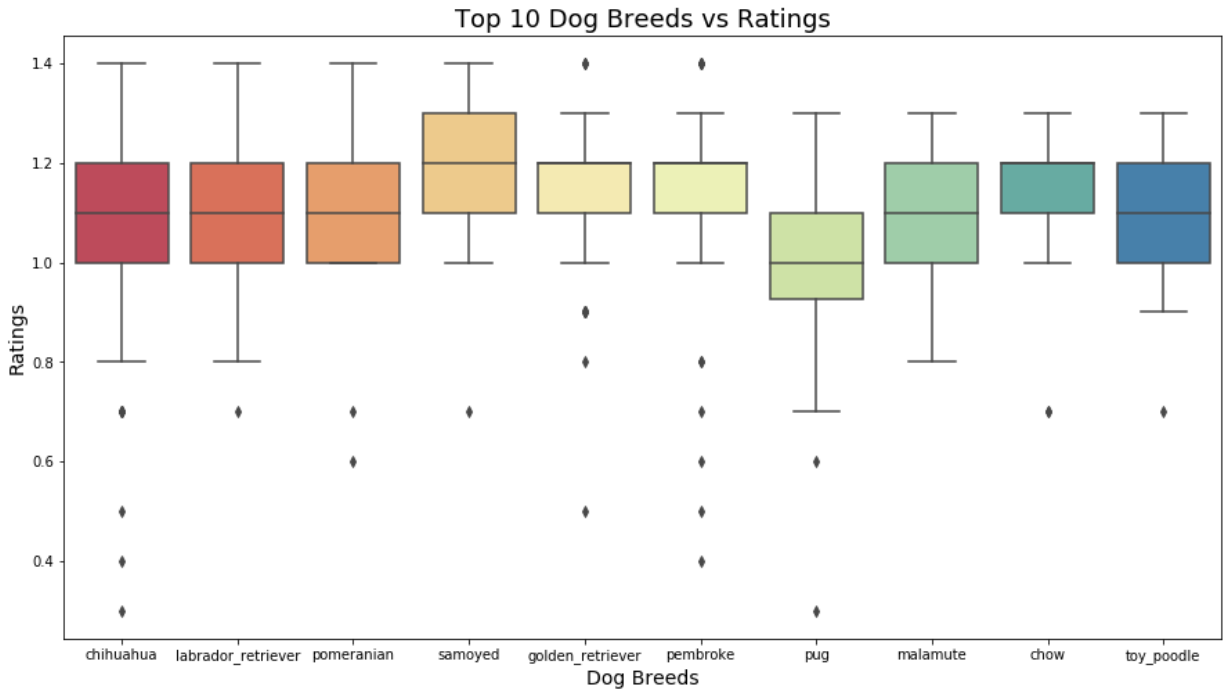
   The image prediction algorithm was able to predict 84.55% of dog images.

   Breakdowns:
   - 1st prediction: 74.07% of tweet images (1447 out of 1994) with 61.44% of mean confidence
   - 2st prediction: 7.37% of tweet images (147 out of 1994) with 11.31% of mean confidence
   - 3st prediction: 3.11% of tweet images (62 out of 1994) with 5.1% of mean confidence
   - no prediction: 15.45% of tweet images (308 out of 1994)

2. Are top breeds of dogs associated with rating?

   Majority of ratings are under 1.40, but there are some extreme ratings (like 177.60 and 42.00). Only 5 tweets have rating more than 1.40. 21% of tweets have rating under 1.00.

   There are 113 dog breeds in the data set. I've used 10 top dog breeds to do the analysis. Samoyed has higher median and higher interquartile range and pug has lower median and lower interquartile range than the rest of top 10 breed dogs. Median rating of all top 10 dog breeds are between 1.0 and 1.2.

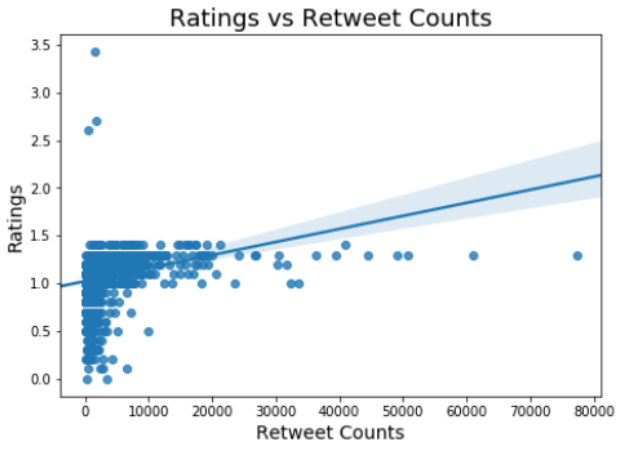


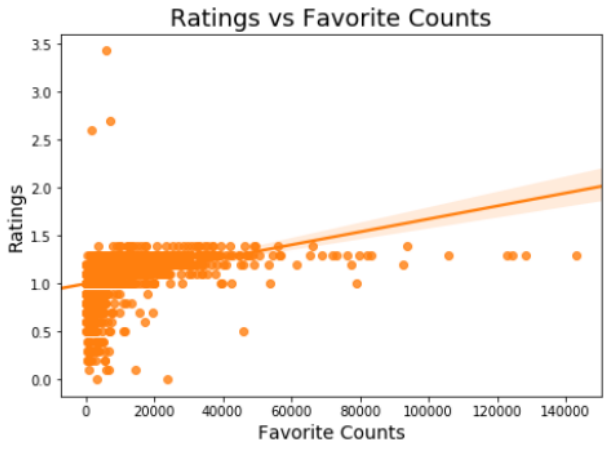
   It doesn't seem like dog breeds have any correlation with rating.

3. Which variables are associated with retweets and favorite counts?

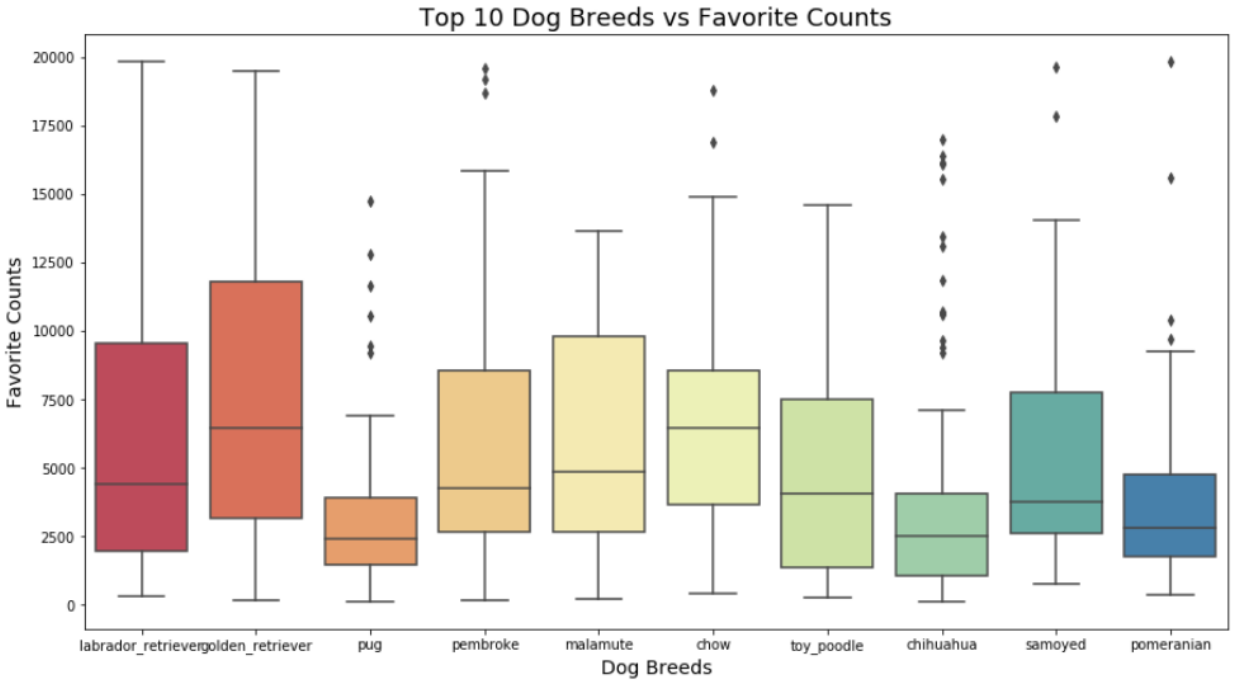
   Analysis using plots:

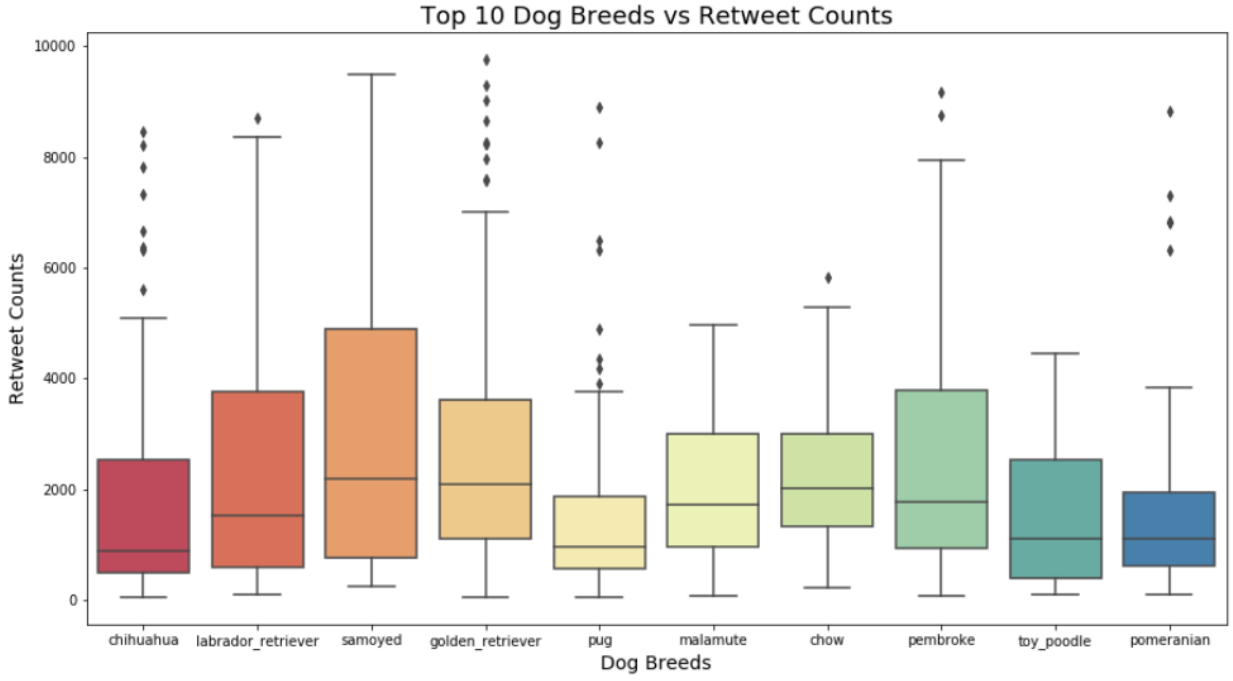
   - For ratings, both retweet and favorite counts have increasing trendline indicating there is positive correlation to ratings.



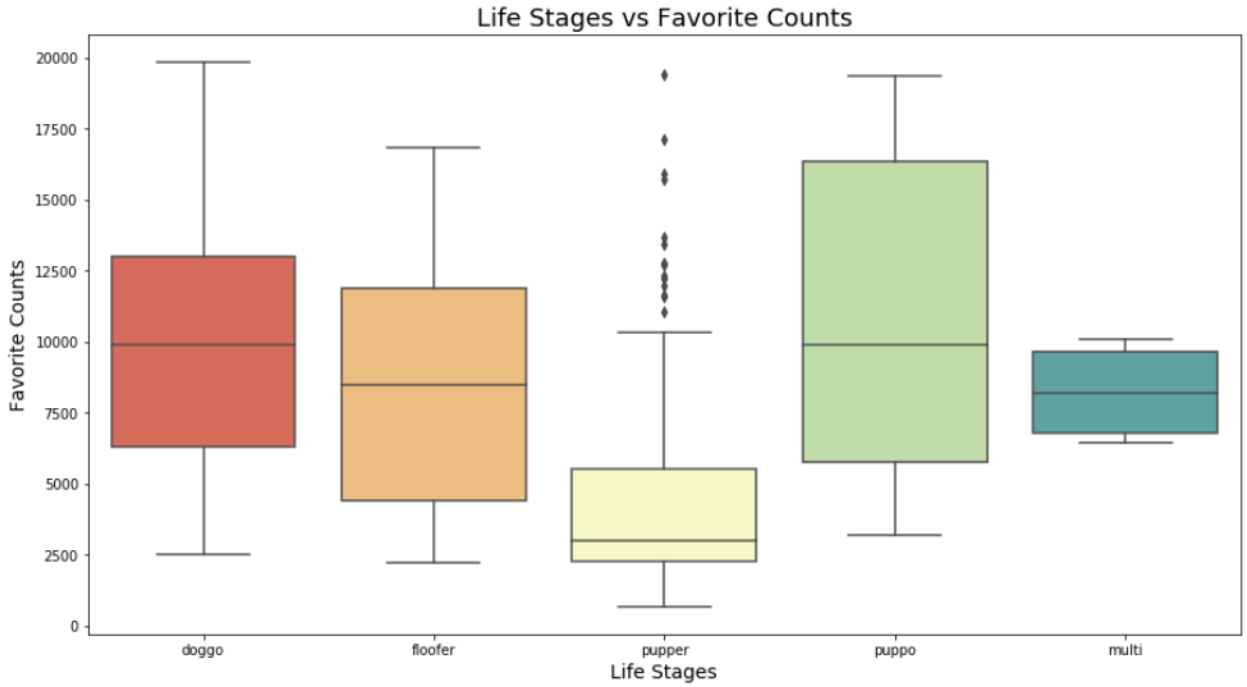


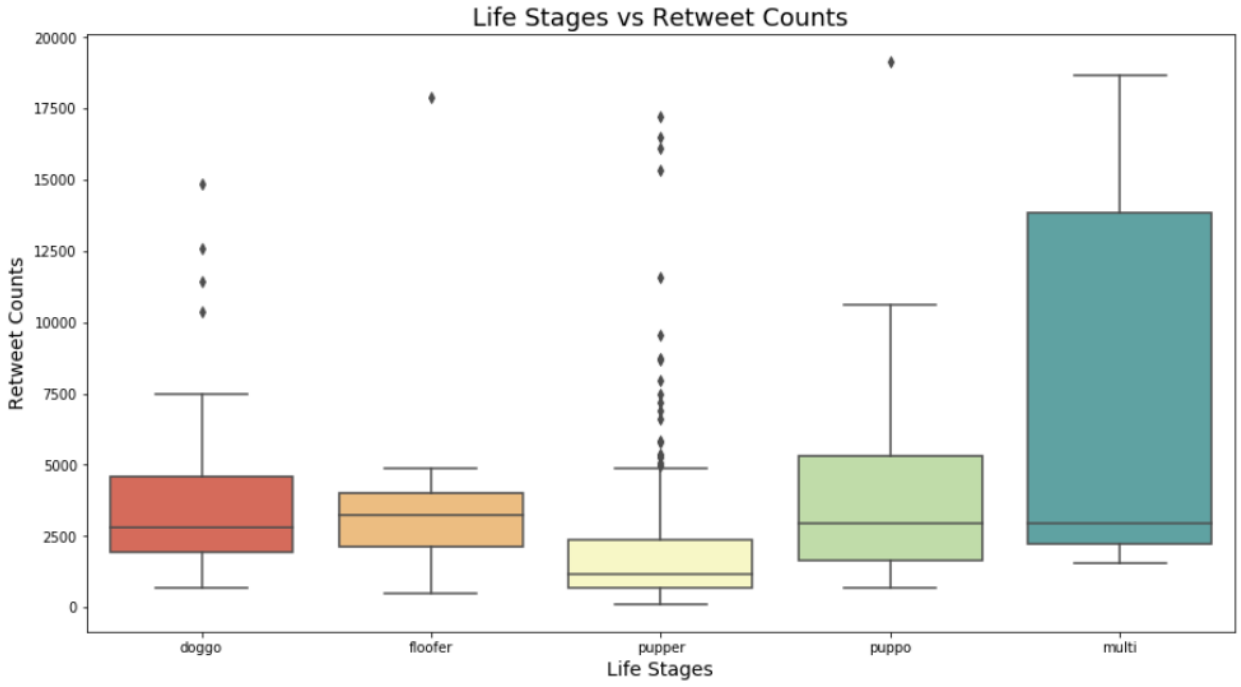
- For dog breeds, I didn't find any significant relationship between dog breeds and retweet and favorite counts.





- For life stages, it seems like pupper gets lower both retweet and favorite counts compare to other life stages. But the difference is not significant. I didn't find any significant relationship.





<u>Analysis using multiple linear regression:</u>

- Rating and life stage 'doggo' have strong correlation with retweet and favorite counts.