

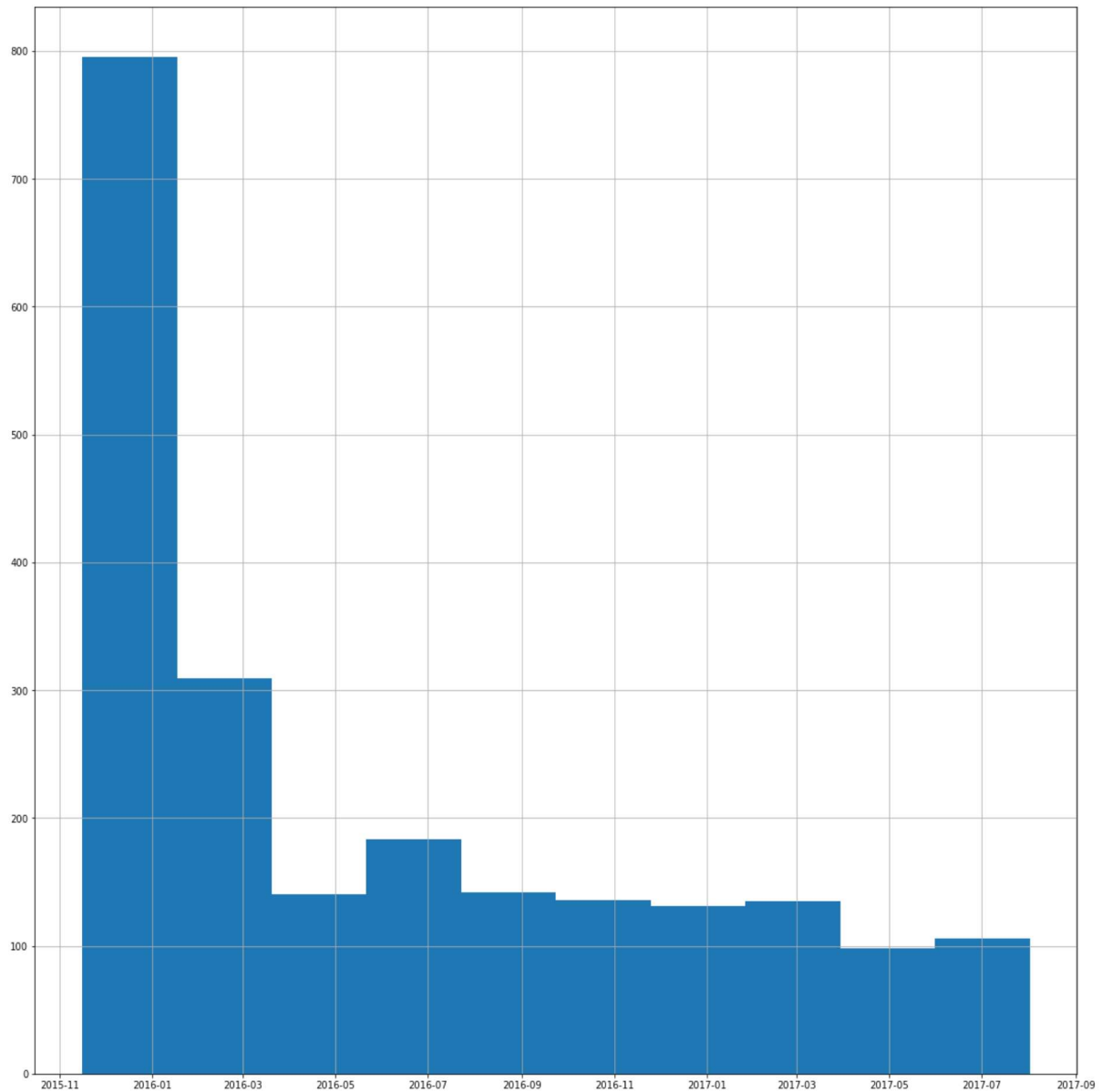
# Report: act\_report

## Insights and Visualisations



After my wrangling efforts i was able to answer the following questions through visualisation:

- What year did the WeRateDogs trend the most on twitter?

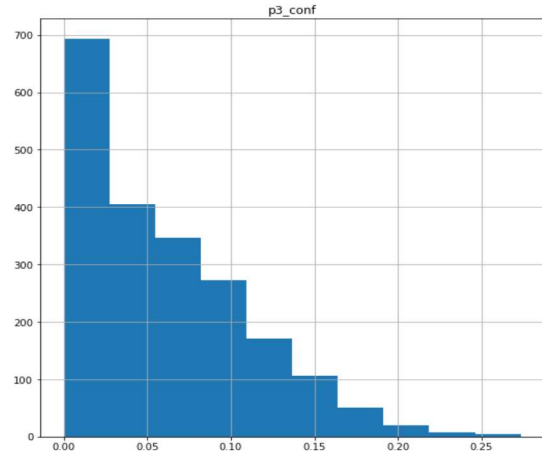
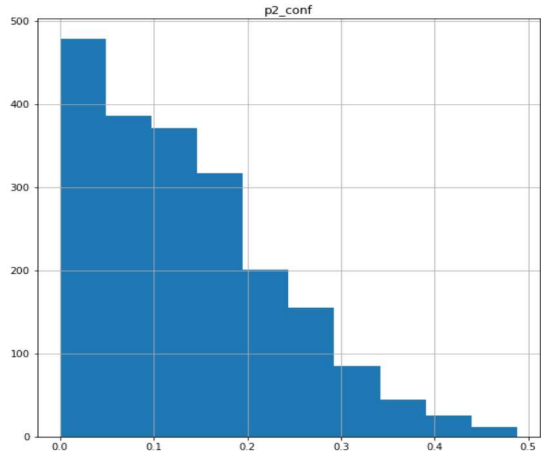
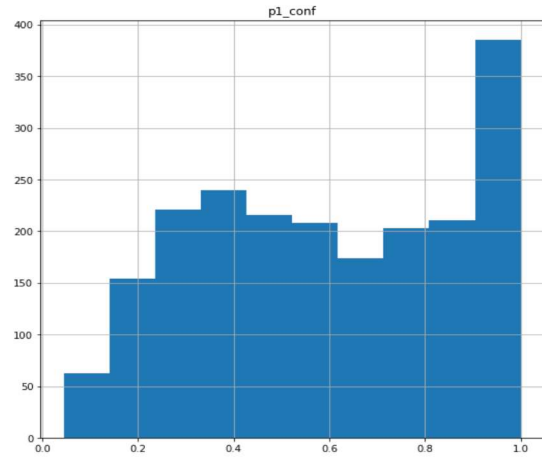
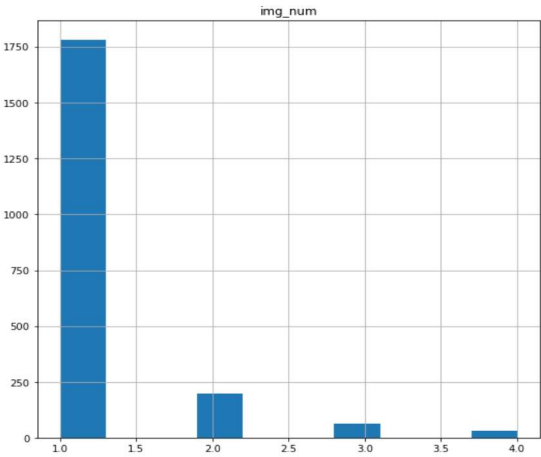


## Insight

the @WerateDogs interactivity, trending and popularity peaked in 2015, and progressively declined over the years till 2017

- which of the trained neural networks performed better?

From the given dataset, each tweet image was run through a neural network to predict the dog breed in those images. Some predictions were correct and some were not. I wanted to get a sense of how correlated the neural network models are in terms of prediction confidence interval. this will enable me understand and know the neural network that will likely make it to production and hence used to classify dogs by names.



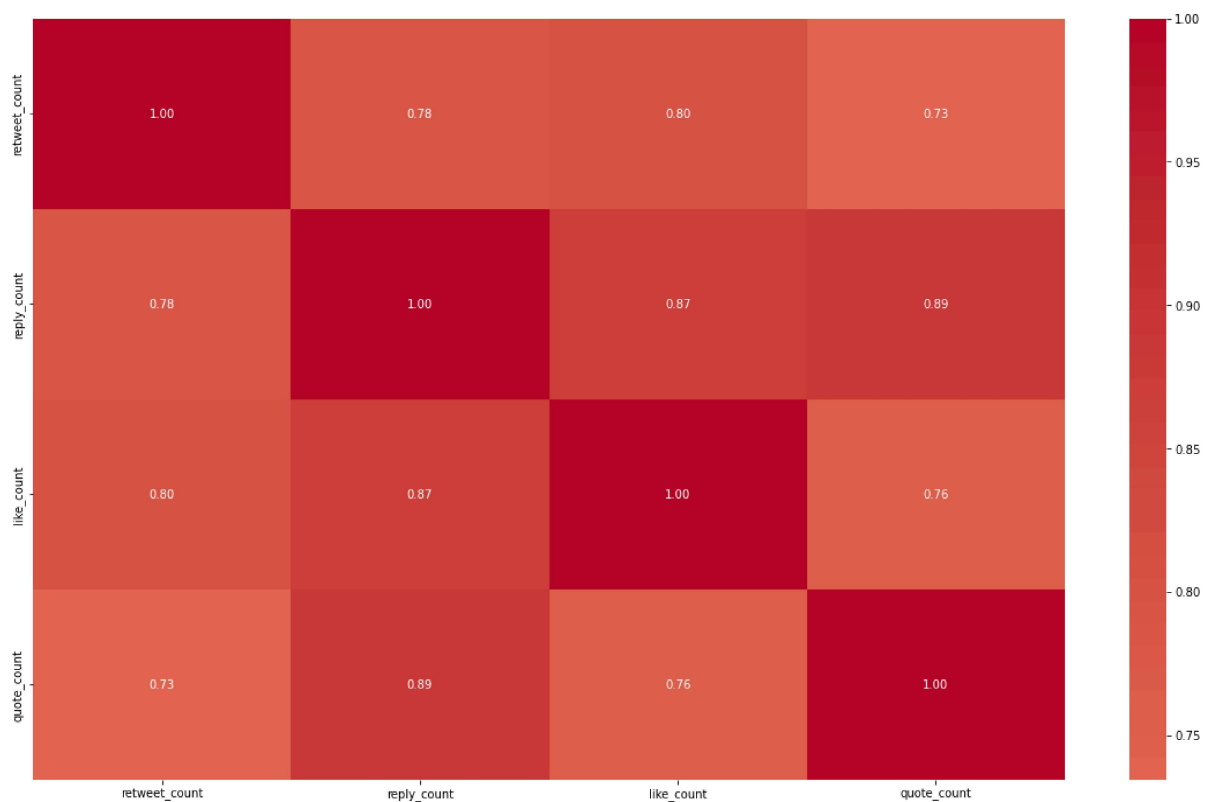
Insights

- looking at the confidence interval distribution, it can be seen that the first neural network with a confidence interval `p1_conf` is more distributed to the right towards 1.0 confidence interval (skewed to the left). so it performs better on average than other trained neural networks on the dataset.

- `p1_conf` and `p3_conf` has high negative correlation(-0.71) that shows a sharp contrast between the two model performances.

- `p1_conf` and `p2_conf` has negative correlation of (-0.51) that still shows how far apart their performances are in opposite directions.

- what are the correlation among the public metrics variables such as `quote_count`, `like_count`, `reply_counts`, `retweet_counts`?



## Insights

there are strong positive correlations amongst quote\_count, like\_count, reply\_counts, retweet\_counts, this depicts that an increase in one leads to an increase in another. This shows that people who likes the tweet are more likely to retweet, reply or quote tweet.

In [ ]: