Act Report

Data Insight

The final data was saved into twitter_archive_master.csv. The data was finally loaded into pandas dataframe. I founded the following insight:

Increase in WeRateDog popularity

It has been evaluated that, the amount of average no. of retweets has increased almost linearly with time. This shows that the popularity of WeRateDog twitter handle increased over time.

Most liked tweet

I found the following tweet as one of the most liked of all, with liked count = 132103.

"*Here's a super supportive puppo participating in the Toronto #WomensMarch today. 13/10*" Link: https://twitter.com/dog_rates/status/822872901745569793/photo/1

Most retweeted tweet

Following tweet was the most retweeted tweet. Retweet count = 78960.

"This is Atticus. He's quite simply America af. 1776/10 https://t.co/GRXwMxLBkh" Link:https://twitter.com/dog_rates/status/749981277374128128/photo/1

Prediction accuracy

I found the total number of tweets, whose dog pictures were not interpreted as a Dog by the neural algorithm. This amount to 504 cases of wrong prediction. overall the accuracy of the algorithm in terms of predicting as Dog was around 74%

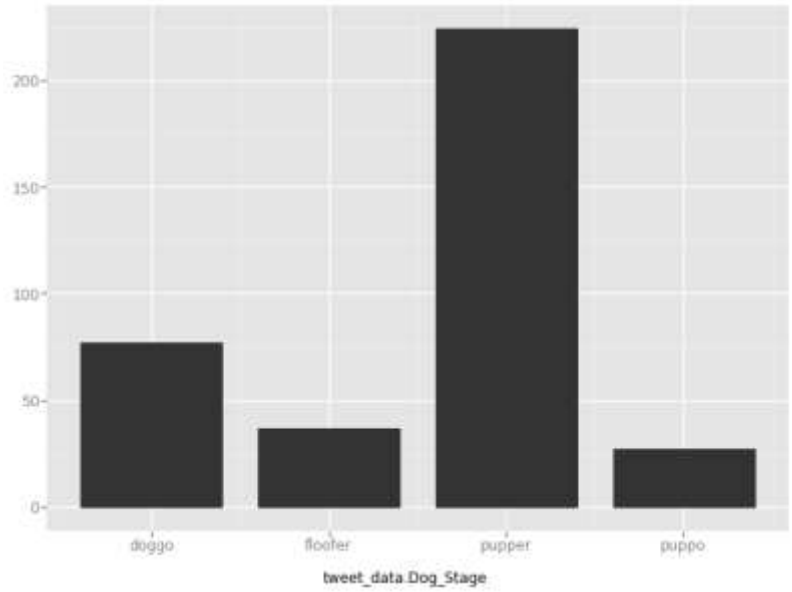
Data Visualization

Data visualization was done using python ggplot libraries. I performed four visual analysis on the final data, as following:

1. Bar Plot for Dog Stage Category

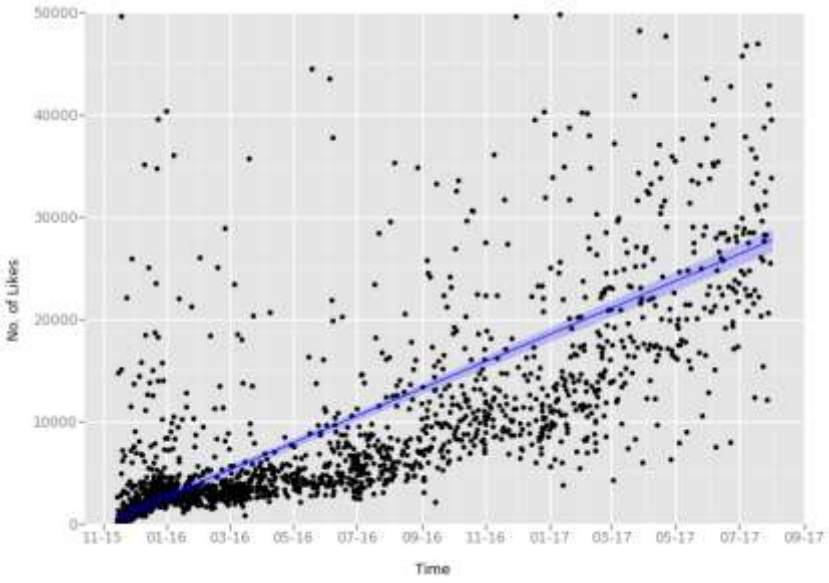
This plot shows the distribution of dog w.r.t its stage. It was found the pupper stage dog were highest

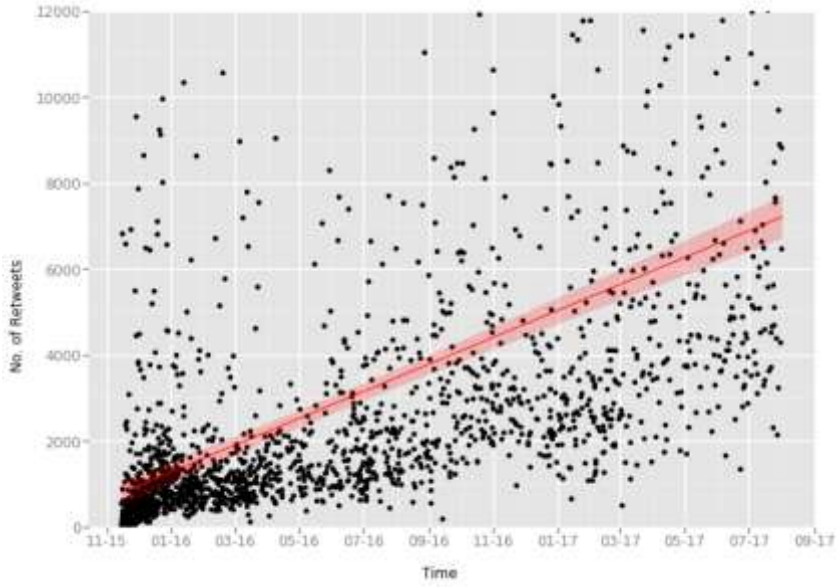
and puppo was least.



2. Trends for twitter popularity in terms of likes received and retweets

it was found and shown in the plot, that with time, the number of likes to a particular tweet and number of retweets increased significantly, which reflected in terms of its popularity.





3. Rating comparison among dogs with different stages

Next, I performed the comparison among the dogs with different stages, in terms which dog type got the highest rating value as per the dog data captured from the tweets.