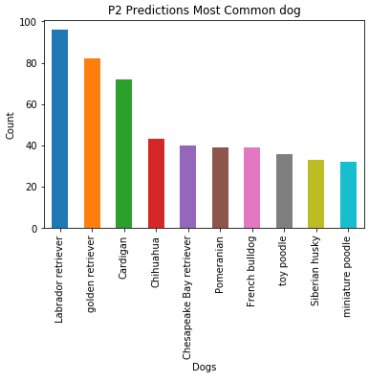
Analysis and Visualization

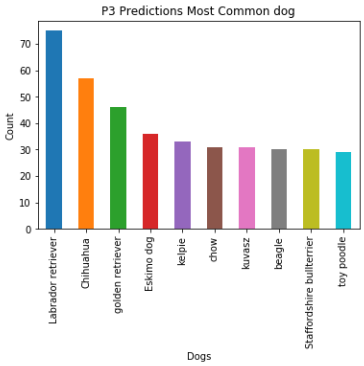
Introduction

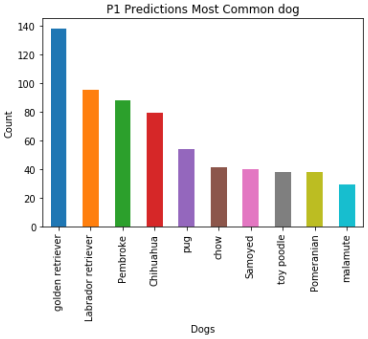
This project is a part of Udacity Data Analyst Nanodegree Where we need to wrangle data from different sources associated with tweets user @dog_rates also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc.

Comparing the predictions of Neural Network

Here we are trying to visualize the predictions given by the neural network with different confidence level. P1 being the most confident and P3 being the least.
Here we can see that for most confidence the Golden retriever was predicted the most and for P2 and P3 Labrador Retriever was the most predicted.



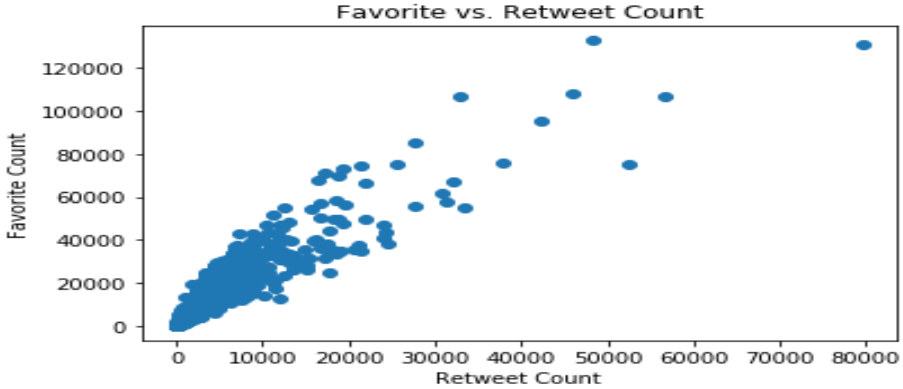


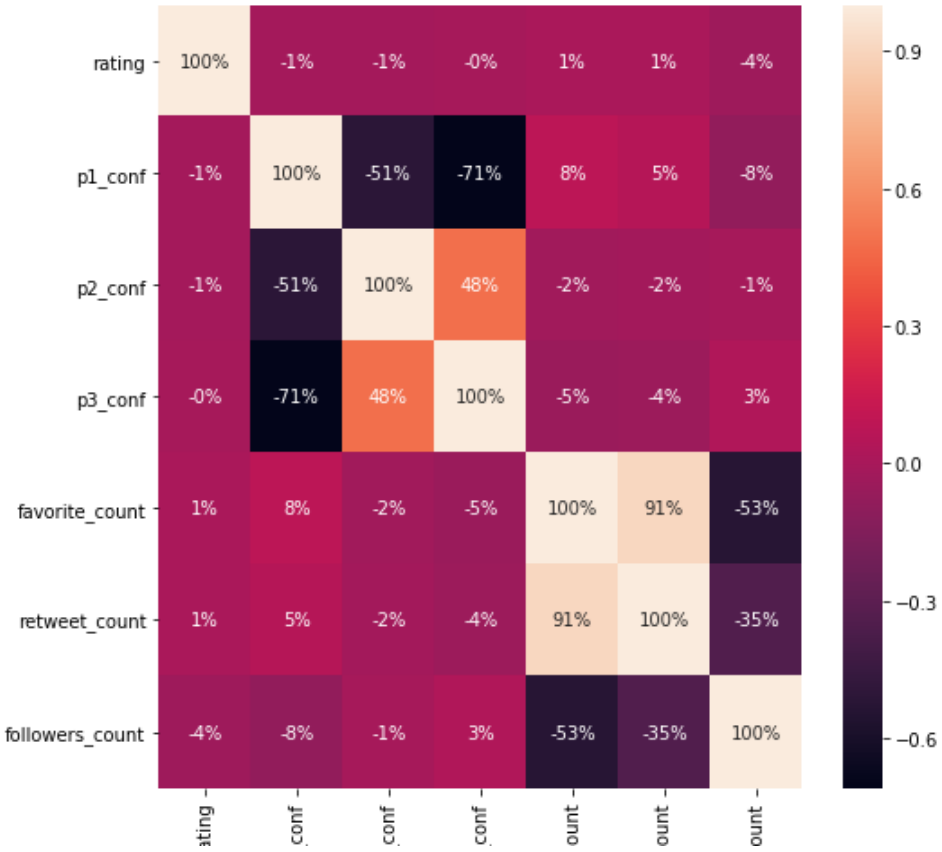


Favourite vs. Retweet Count

Here I thought of getting a heat map for all the correlation and check which one have the highest correlation and found out that Favourite Count and Retweet Count have a positive correlation between each other.

Here we can see that when one is increasing the other is also increasing and all the major retweet and favourite lie between 70000 favourites and 25000 retweets. The most popular tweet has about 130000 favourites and 80000 retweets.





Most Common Dog Stage and Source Type

Here we tried to gather all the dog stages which are present from the obtained list and find out the most popular one is Pupper.

Similarly, the most used source is iPhone.