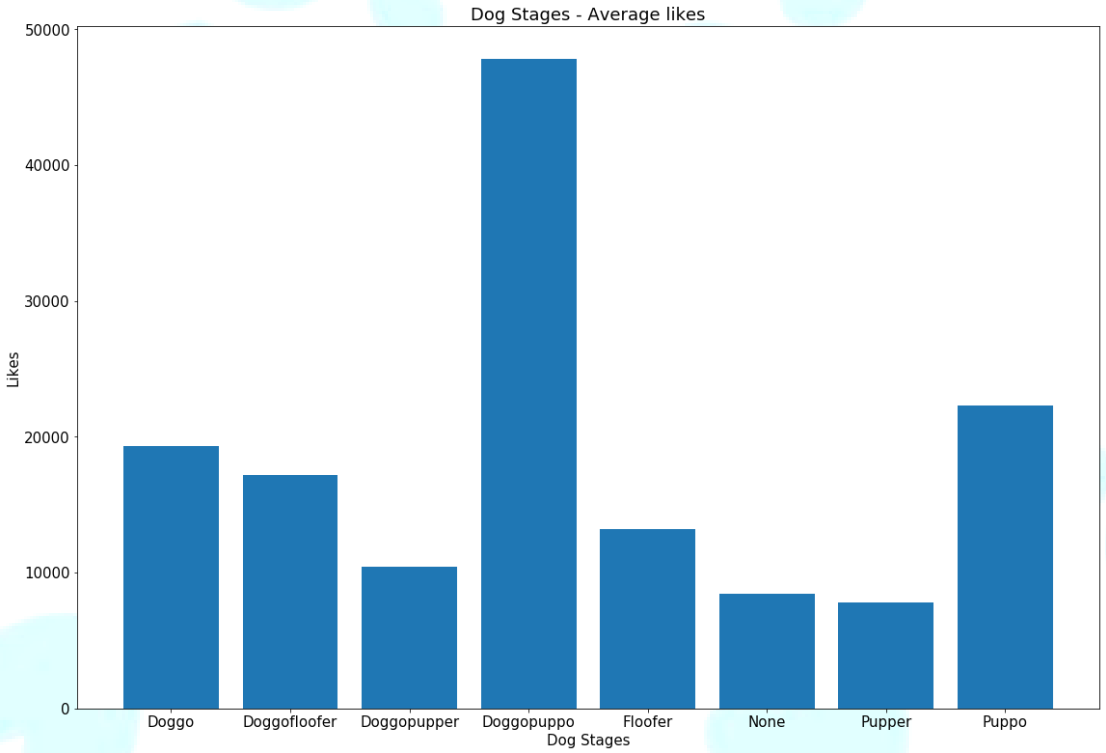
Act Report

This document communicates all the insights and displays the visualization produced in the **wrangle_act** document.

After gathering, assessing and cleaning the WeRateDogs twitter data, the master dataset was saved to a CSV file named "twitter_archive_master.csv". I analysed this dataset to establish some insights from the data and generate the visualization below.



The tweets that categorized dogs as '**doggo and puppo**' received the most likes on an average basis more than the other dog types followed by **puppos** and **doggos** individually with **pupper** being the least desirable dog type. This finding is fully supported by the fact that the highest retweeted tweet belonged to a **doggo** while the most favorited tweet belonged to the **puppo** dog type.

(Most favorited tweet)

A strong positive correlation was observed between ***retweet_count*** and ***favorite_count*** with a correlation coefficient of **0.9**. The more a tweet was retweeted, the more it was liked/favorited.

From the cleaned data, it could be safely assumed that the neural network responsible for the predictions of the dog types was very effective. This is evident in the values in the p1, p2 and p3 columns. On the average, the p1 values were higher than the subsequent predictions in p2 and p3, their mean values also followed this pattern. If this wasn't the case, it would have resulted in an anomaly as the probability of the predicted values should reduce with each successive prediction.

Out of all the predicted dog types, the **golden retriever** emerged as the one with the most predictions as it had 135 out of the 1453 predictions that were made, it was followed by the **labrador retriever** with 91 predictions. The least predicted dog types were the **scotch_terrier, entlebucher, groenendael japanese spaniel, clumber, standard_schnauzer and silky terrier** , with each one having 1 prediction.