Project – Data Wrangling

Insights of the Analysis of We Rate Dogs Twitter Data

The analysis and visualizations will be based on the data that has been wrangled (gathered, assessed and cleaned) in the notebook, 'wrangle_act.ipynb'

Insight 1:

These insights are based upon the number of tweets , percentage of tweets and favourite count and retweet count

- Number of tweets with rating 10+: 1665
- Percentage of tweets: 80%
- Tweets having favorite count <= retweet count: 0

Insight 2:

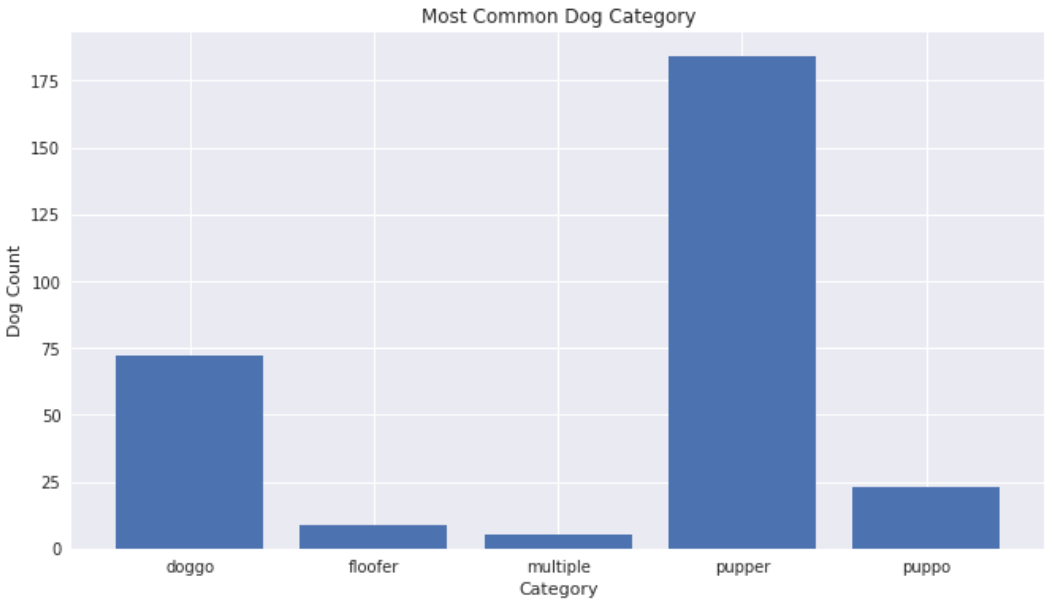
These insights are based upon the retweeted tweets and favourite tweets, the data which we got from the Twitter API.

- Top 5 retweeted tweets and top 5 most favorite tweets:
  Given in the 'wrangle_act.ipynb' notebook

Insight 3:

This insight tells us the most common Dog Breed

- Most Common Dog Category



As we can see, Pupper is the most common dog category.

Other Insights which have been drawn after the analysis:

- Most popular dog type is a pupper
- The most popular dog, based on image predictions, is a Golden Retriever
- The median and mean retweet count is 1226 and 2590.42 respectively
- The median and mean favorite count is 3529 and 7999.35 respectively
- The median and mean ratings out of ten are 11 and 11.69 respectively

Before drawing the insights, the separately cleaned data sets have been combined into one master data set and post that all the analysis has been done.

The insights have been drawn on the data that has been gathered, assessed and then cleaned in accordance with the requirements as specified by Udacity.

These insights are entirely based on the cleaned master data, ‘twitter_archive_master.csv’