ACT_REPORT - WE RATE DOGS – TWITTER DATA

1.0 Introduction:

This article communicates the insights of my finding generated from the analysis of the WeRateDogs Twitter datasets. This project contains 3 datasets, in which 2 of the datasets were merged together as a result of having similar tweet_id. However, I'll briefly talk about my analysis process and also communicate the results and in conclusion display visualizations. I started from retrieving the dataset from the server using a pandas function, then assessed and ran an analysis and visualized the data.

Gained insight into 3 aspects of the data, namely: Characteristics of Dogs with high rating, what source was utilized the most, most Popular dog breed. Then went ahead to visualize one insight.

2.0 Stored Data

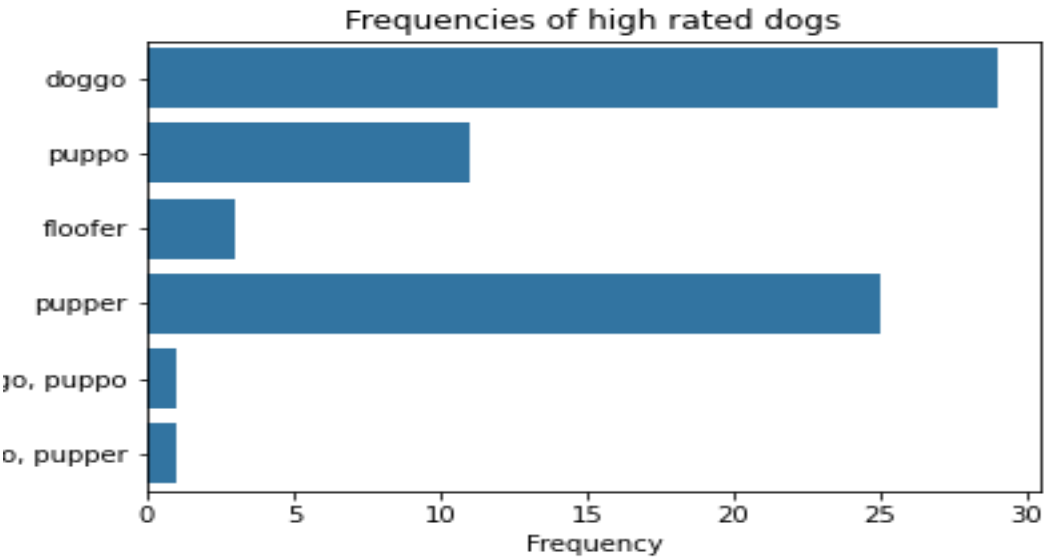
In this section, I stored the gathered, assessed, and cleaned twitter_archive_image, retweet_fav_count dataset to a CSV file named "twitter_archive_master.csv", "retweet_count.csv" respectively. Then I ran the saved datasets on the notebook, to view the data.

3.0 Analyzing and Visualizing Data

In this section, I analyzed and visualized my wrangled data. In which I developed **three (3) insights and one (1) visualization.**

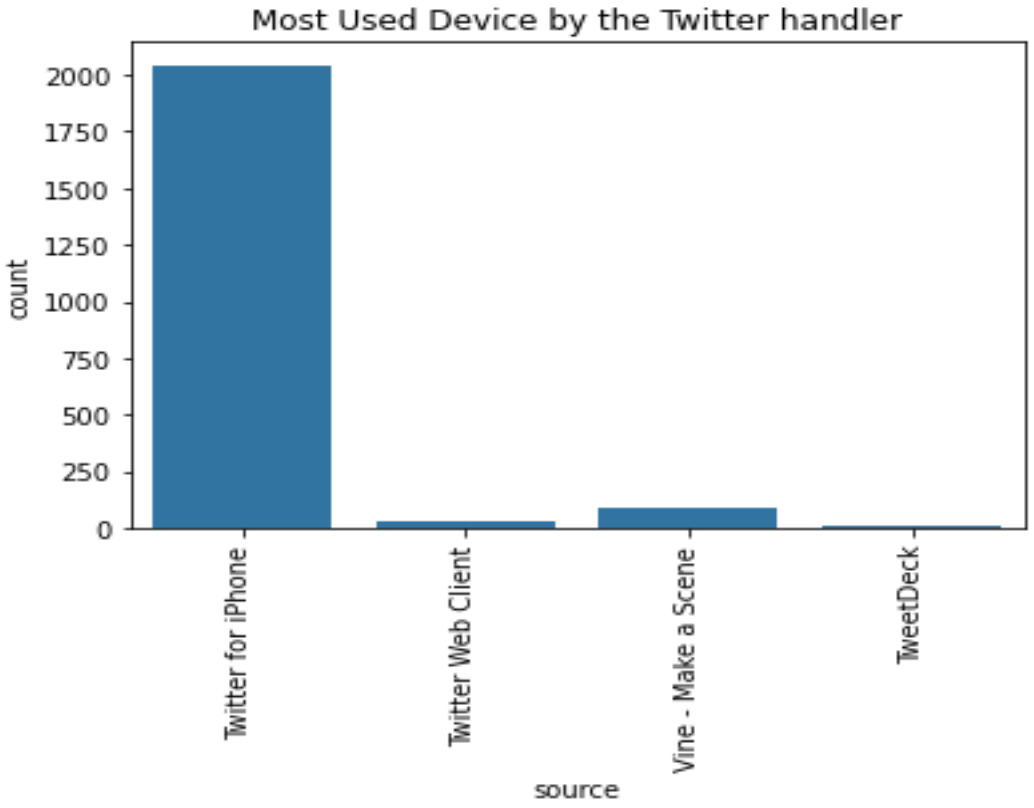
3.1.1 Characteristics of Dogs with high rating

In this section, I looked at the characteristics of Dogs with high rating from 13 upward. It was observed that Doggo type of dogs were much in this category, with 29 counts. Thus, it could be deduced that Doggo dog's Stage were most tweeted in the WeRateDogs Twitter data.



3.1.2 What source was utilized the most. - Twitter for iPhone

In this section of the analysis, I looked at the source that was more utilized in tweeting the dog posts. I also visualized my finding. From the result, it could be deduced that, iPhone had more counts, thus iPhone was more utilized in the process. iPhone had a count of 1506.

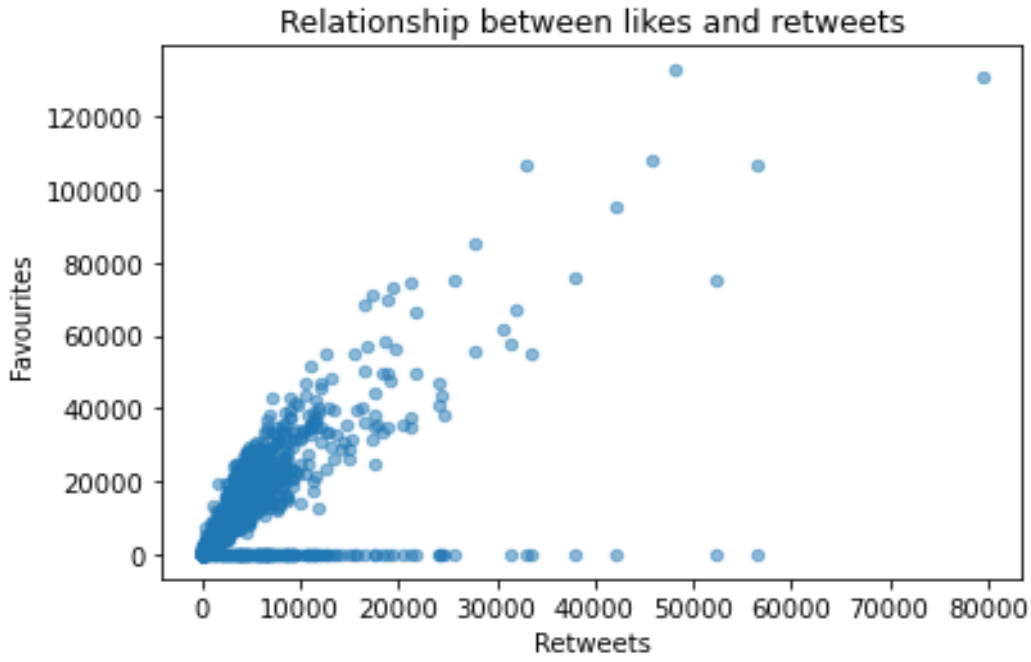


3.1.3 Most Popular dog breed - Golden Retriever

In this section of the analysis, I determined the most popular dog breed, in which I Used value counts to determine the number of times a particular dog breed occurred, Golden Retriever had more occurrence. - 139.

4.0 Visualizations:

In the visualization aspect, I used bivariate plotting to display the relationship between retweet and favorite(likes) using pythons' package - maplotlibs scatterplot.
However, the result displayed from my plot showed that the magnitude of Retweets by the WeRateDogs twitter handler are positively correlated with favorites (the amount of time the tweet was liked by twitter users.


Relationship between likes and retweets
Favourites
Retweets
0
20000
40000
60000
80000
100000
120000
0
10000
20000
30000
40000
50000
60000
70000
80000