Act-Report

Introduction

WeRateDogs's Twitter account have over 4 million followers that rates people's dogs with a humorous comment about the dog. I try to establish where and what drives WeRateDogs's Twitter Activity.

To fulfil the requirements for the wrangle and analyze course on the Udacity Nanodegree program, the first three questions are reported in this submit.

- Top 5 & Bottom 5 Counts of Most Popular dog breed
- Top 5 & Bottom 5 Counts of Most Popular dog stage
- Source of WeRateDogs Twitter Traffic

Additionally, I will try and explore the relationship between the top dog breed/stage likes and retweet counts and compare with the bottom dog breed/stage. However, that will not be part of this report.

Note: I assumed that there is a relationship amongst the breed of dogs, stage and the level of traffic on the WeRateDogs's Twitter account.

Source of WeRateDogs Twitter Traffic

Using a simple visualization.

```
{twitter_data_clean.source.value_counts(normalize=True).plot(kind='bar')}
```

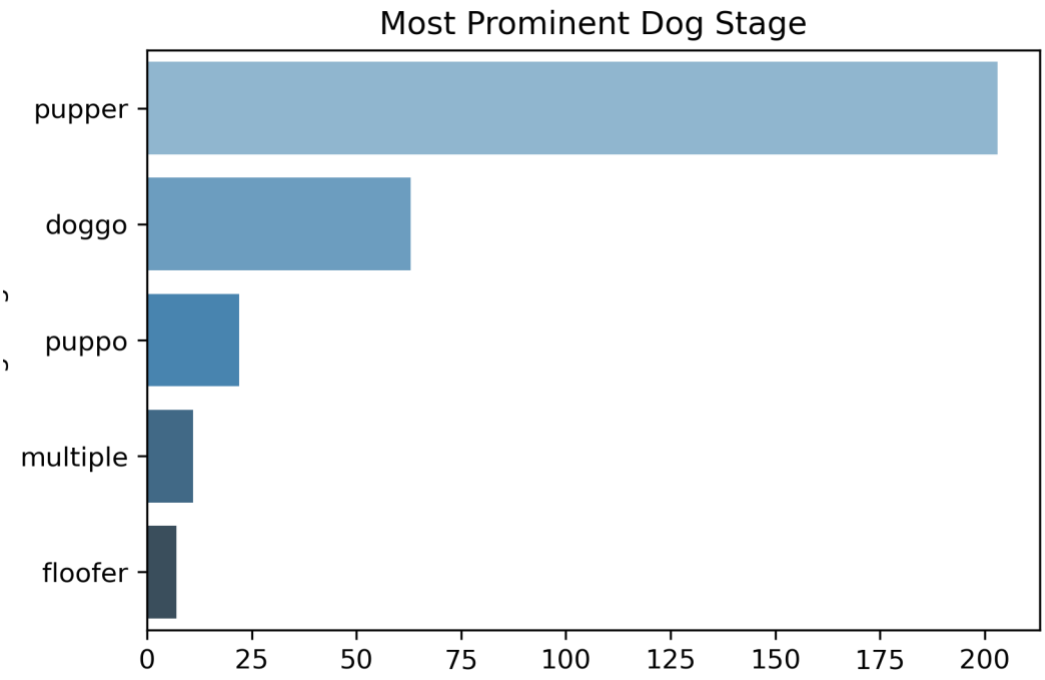
I was able to visualize the frequency distribution of the sources of **WeRateDogs** Twitter Traffic. 'Twitter for iphone' brought over 80% of the traffic and remains the most relevant source of traffic to the account.

Top 5 and Bottom 5 Popular Dog Breed

Using the predictions columns, I filtered out the predictions that are not dog breeds and returned the top and bottom 5.

Top 5 and Bottom 5 Popular Dog Stage

During the data cleaning stage, I combined the four columns into one called stage and I did a value_counts that returned the top and bottom 5.


Most Prominent Dog Stage
pupper
doggo
puppo
multiple
floofer
0
25
50
75
100
125
150
175
200