Udacity Data Analyst Nanodegree

Wrangle and Analyze Data Project

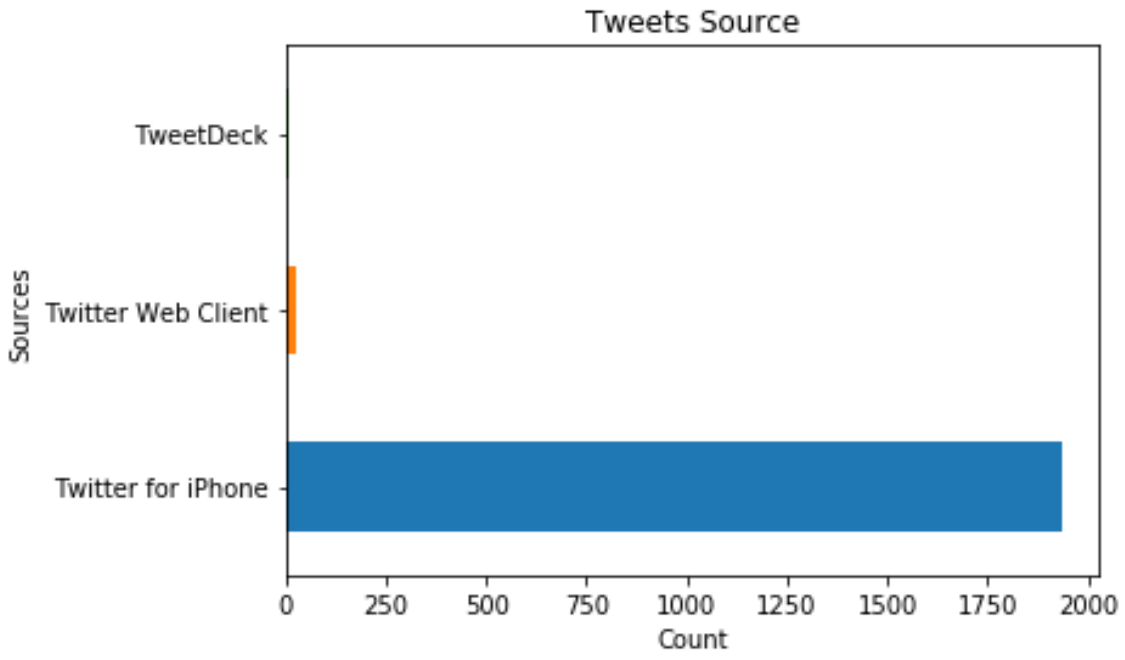
Analysis of WeRateDogs™

Introduction:

The dataset that you will be wrangling (and analyzing and visualizing) is the tweet archive of Twitter 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. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

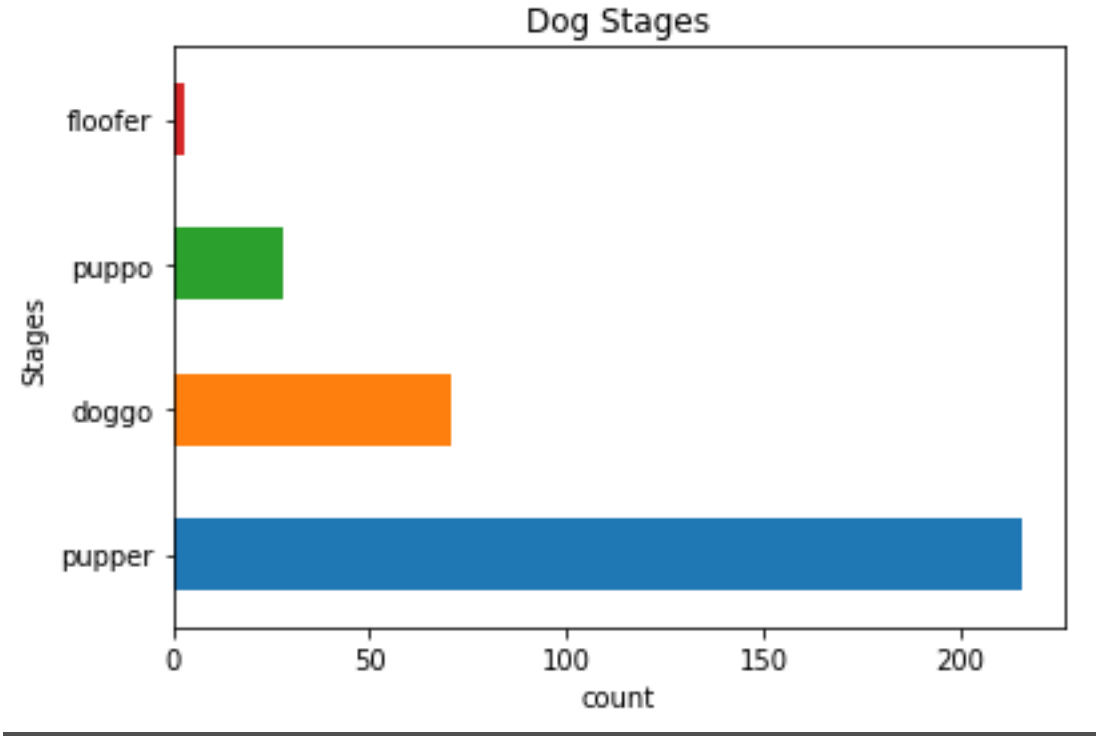
Insights and displays the visualizations:

Tweets' Source:



It is clear that the iPhone as Twitter's source is extraordinarily pretty much exactly 98%.

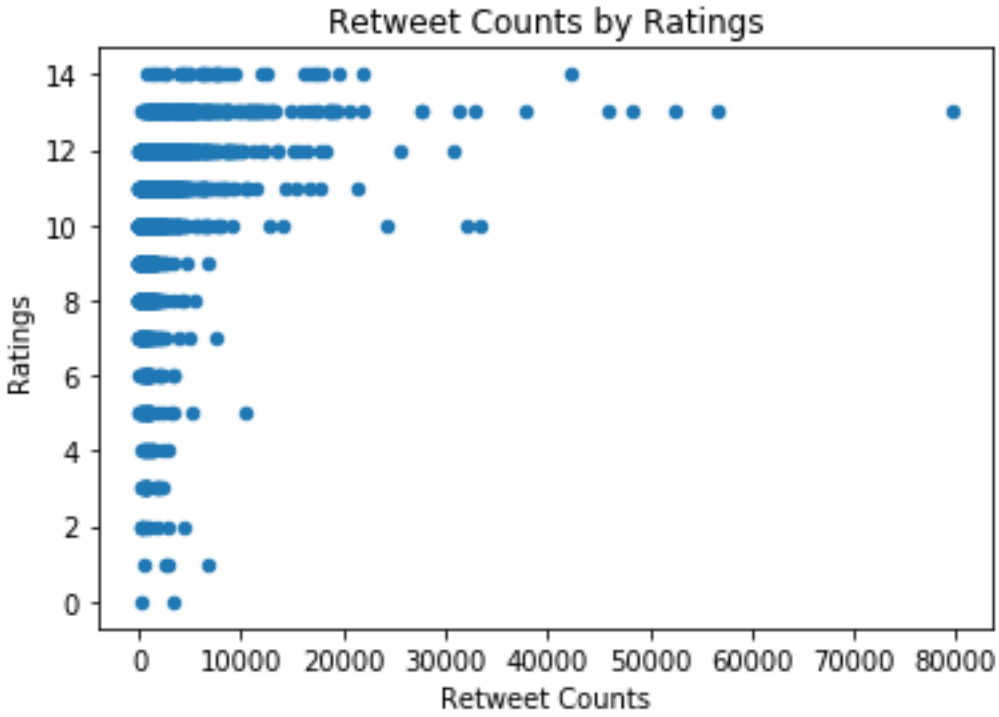
Dog stages:



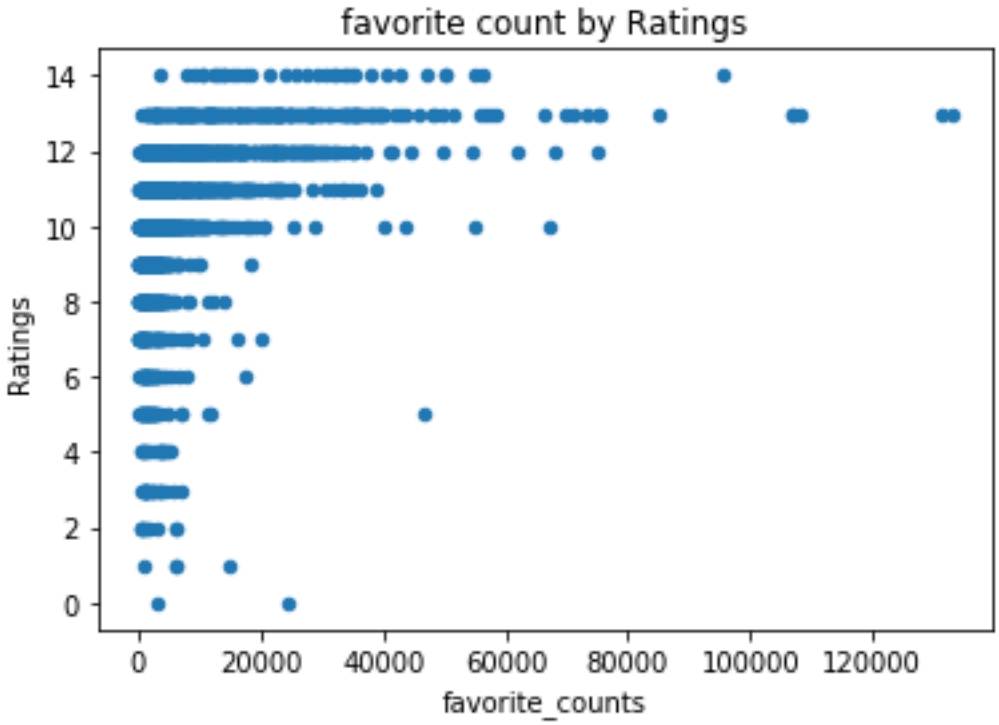
The above graph shows that the most popular Dog stage is pupper and the least one is floofer.

Retweet and favorites Counts vs. rating:

I would like to know if the amount of retweets and favorites is related to the rating from WeRateDogs:



The above graph shows that the highest ratings do not receive the most retweets.



The above graph shows that the highest ratings do not have the highest favorite counts but I found that a statistically significant positive effect from WeRateDogs ratings on the number of retweets and favorites.

The retweets count and favorites have the same distribution, but the favorite counts is much more that retweets count,

Max Retweet: 79515

Max favorite: 132810

The tweet with highest retweets count has 131075 favorites, and the one with highest favorite counts have just 48265 retweets.

This makes sense since more people will favorite a tweet instead of taking ownership of a retweet, so this information wasn't unexpected.

Thank you…