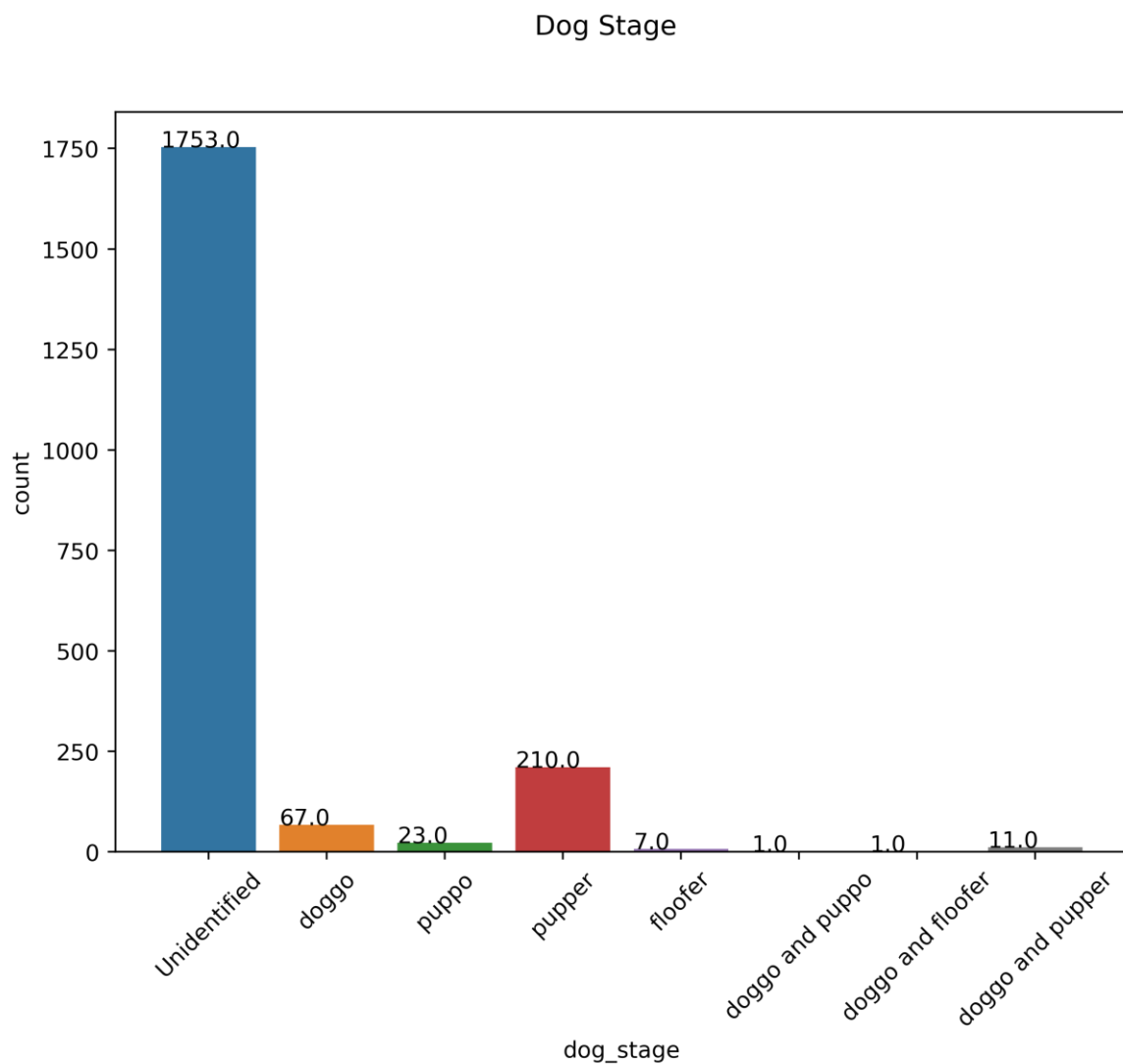# Report: act_report

From merged cleaned Dataset, some few visualization were made based on the data available and also was able to tackle some questions:
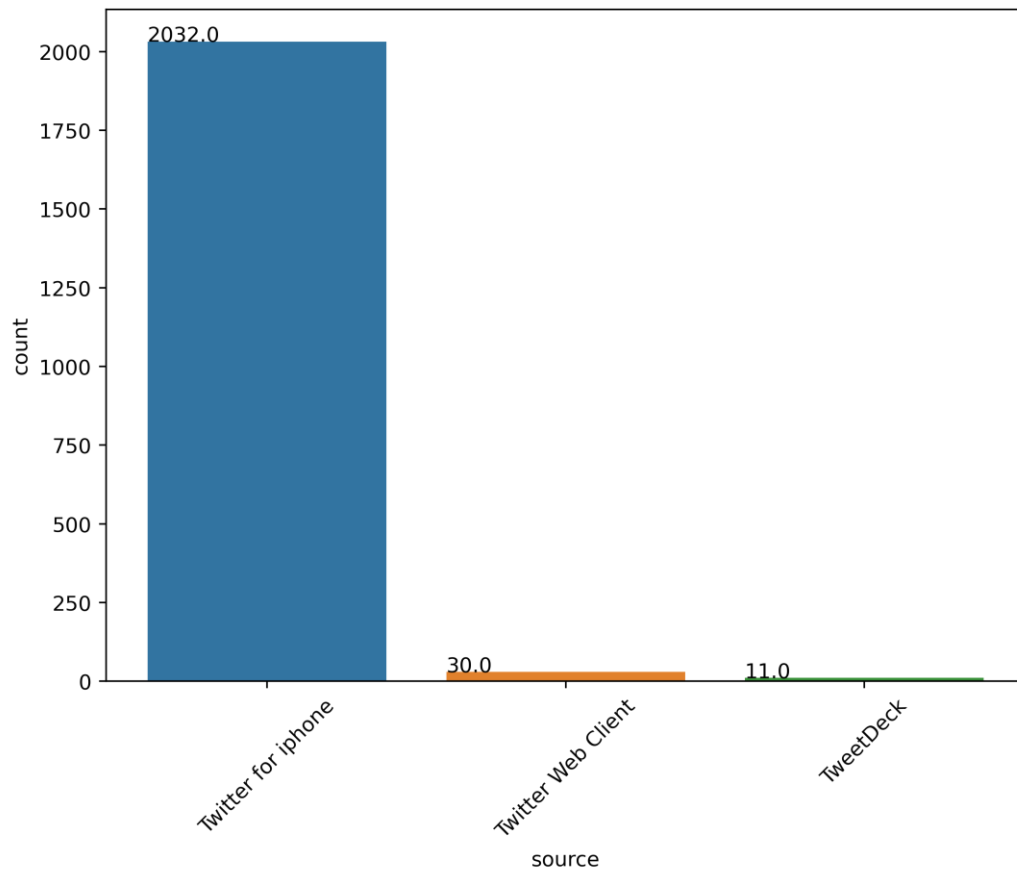
Visualization 1:

Dog Stage

Observation: Most of the observations in the Dog stage category have no dog stage attribute, so they have been classified as unidentified. Out of those that have dog stage attribute, floofer dog stage has the lowest occurrence with 7, followed by puppo with 23.Pupper dog stage is ranked highest with 210, and doggo is ranked second with 67.
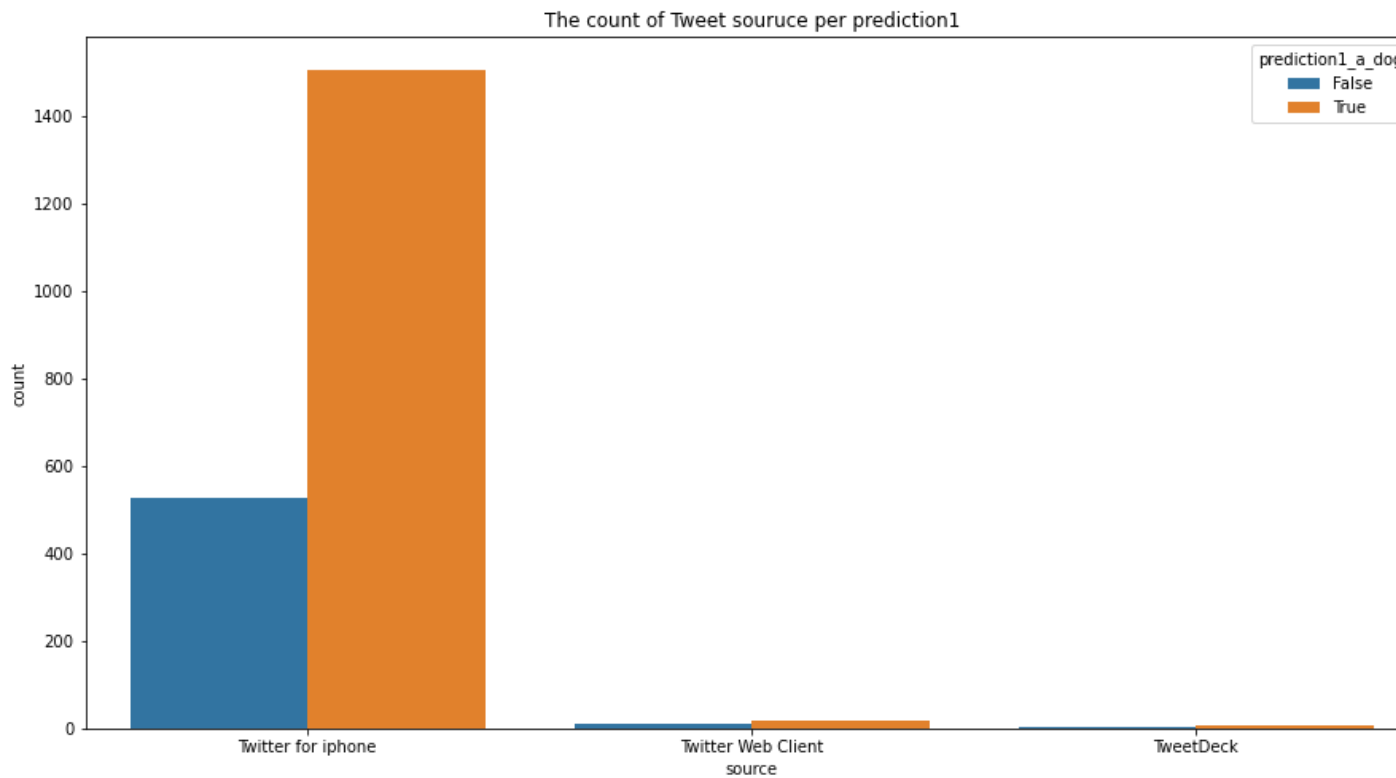
**Visualization** 2:

Source of Tweet Distribution

From the above, we are also able to determine user's brand of mobile phone. More Iphone users observed to engage in the tweet with their phone.

Visualization 3:

The above shows the disparity between the predictions of each of the tweet source. Though many tweet came off iphone, most of the predictions on were true as there is a huge difference between the true prediction and the false prediction, but the difference from other source are at pal.

Visualization 4:

First Prediction Validity by Twitter Retweet Correlation
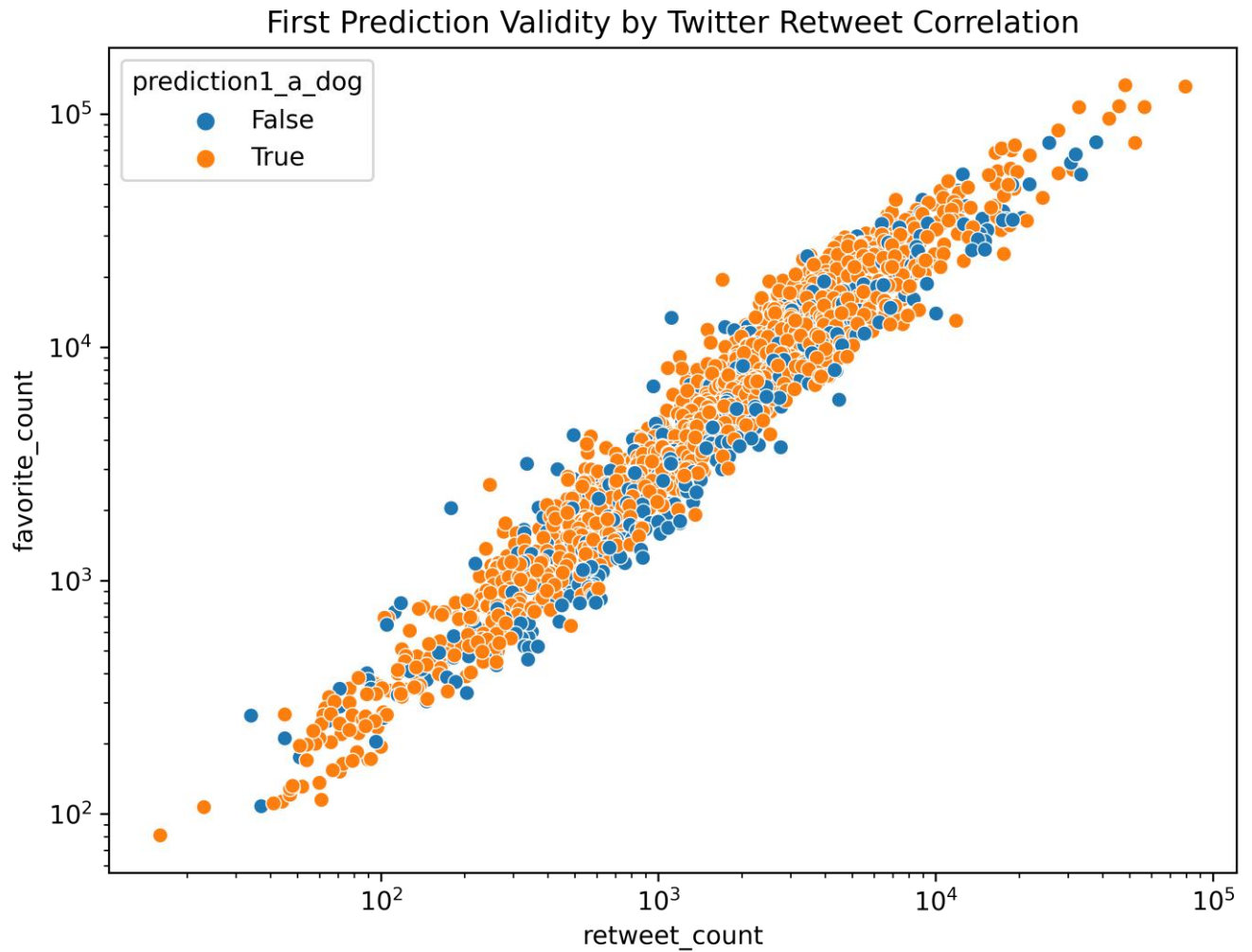
Lastly, Since Prediction1 has better quality, there will be limit of analysis to prediction1. However, Retweet count and favourite count correlate positively, it seems logical that the more a tweet is retweeted, the more likelihood it has for gaining more followers for potential likes. Data points are categorical by the validity of the second prediction, the correct prediction is true and the incorrect prediction is false.