# Act Report

## Overview

The Analysis and Visualizations were focused on the percentages of tweets from different sources, the most common dog stages, the most used dog names, and correlation between "favorite_count" and "retweet_count".

## 1. Tweet Sources

From the data acquired, tweets of this account came from three sources; iPhone,TweetDeck and Twitter web client. Out of the 1986 tweets, 1,948 tweets (i.e 98.09%) came from iPhone sources.
The second-highest percentage of tweets—1.41%—came from the Twitter Web Client.
The smallest percentage of the tweets, 0.50%, came from Tweets via TweetDeck.
The percentage distribution of each tweet source is displayed in the visualization below.
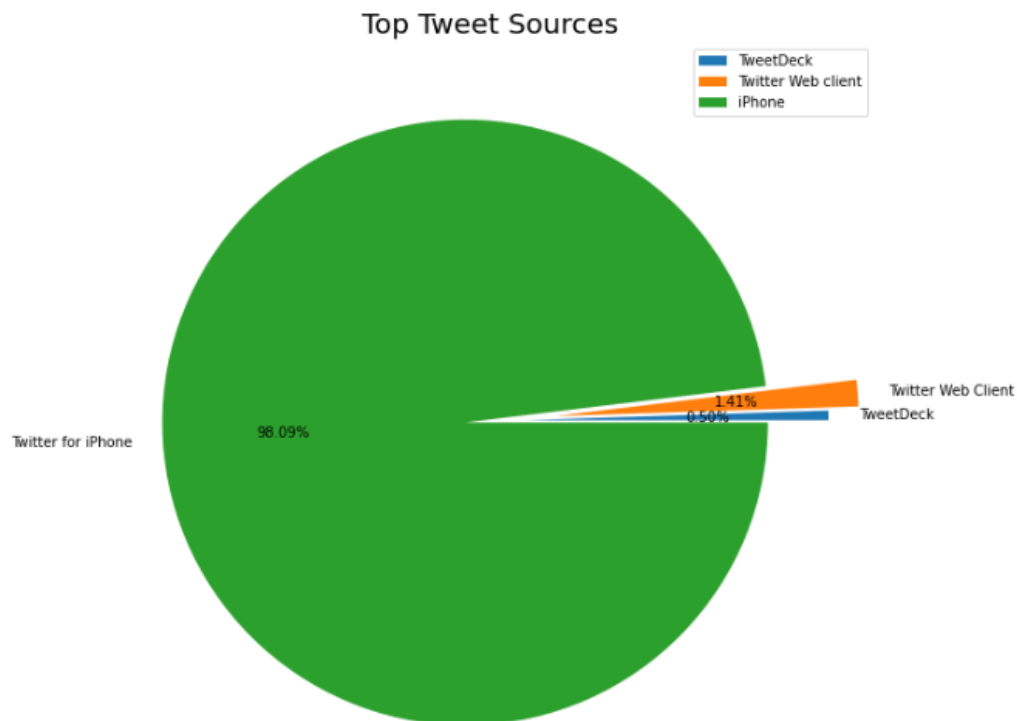
## 2. Dog Stage Ratio

Though not for all tweets, the data provided us with the stage of life of some dogs.
The main stages of the dogs provided were *doggo, pupper, puppo, and floofer.*
Yet there were some few dogs that were classified under two stages.
Our visualization shows that the *pupper* stage is the most common, followed by *doggo*, then *puppo*, then *pupper,doggo* stage, then *floofer*, then *puppo,doggo* and lastly *floofer,doggo* as the least common dog stage in the dataset.
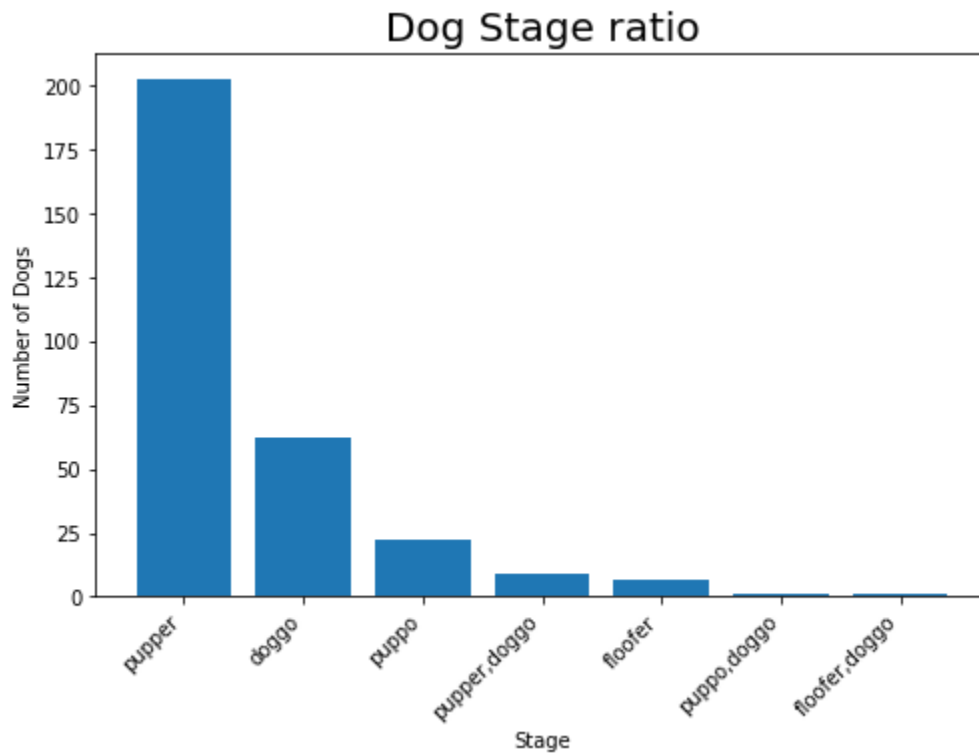


## 3. Most Used Dog Names

According to the data, the most popular dog names are 'Oliver', 'Cooper', and 'Charlie'.
The visualization below also displays the top 10 most popular dog names used.

Most Used Dog Names
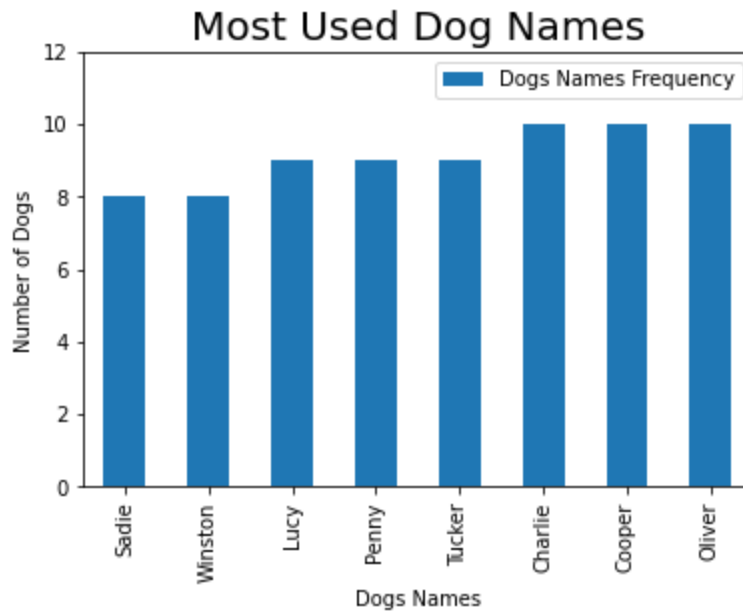
## 4. Favorite Count Vs Retweet Count

The "favorite_count" column within the dataset provides the number of times the tweet has been favorited.

Similar to the "favorite_count" column, the "retweet_count" column shows how many times a tweet has been retweeted.

The visualizations below show the distribution of retweet_count, distribution of favorite_counts, and also the correlation between the two counts ( retweet_count vs. favorite_counts ).

Observing the histogram plot for the distribution of favorite_count and that of the distribution of retweet_count, it can be seen that both graphs are positively skewed to the right.

The scatter plot of values from both columns shows a very strong positive correlation between favorite count and retweet count with the correlation 0.92

Distribution of favorite_count


Distribution of retweet_count

favorite_count vs. Retweet_count