Data Wrangling Project: WeRateDogs

The following report provides three insights as a result of analyzing and visualizing the data provided in this project.

Introduction

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. The account was started in 2015 by college student Matt Nelson, and has received international media attention both for its popularity and for the attention drawn to social media copyright law when it was suspended by Twitter for breaking these aforementioned laws.

The data sources:

- Enhanced Twitter Archive file.
- Image Prediction file.
- Additional Data via the Twitter API.

Data Analysis and Visualizations insights:

- What is the most common dog stage?
- The distribution of tweets sources.
- Relationship between favorite and retweet counts.

1- The most common dog:

The figure shows that the 'pupper' is the most common dog stage followed by 'doggo' and 'puppo'. This may be due to his size and characteristics.
It was expected, as per Google, small size dogs adapt well to apartment living, affectionate with family, kids friendly, easy to groom, easy to train, and friendly towards strangers.
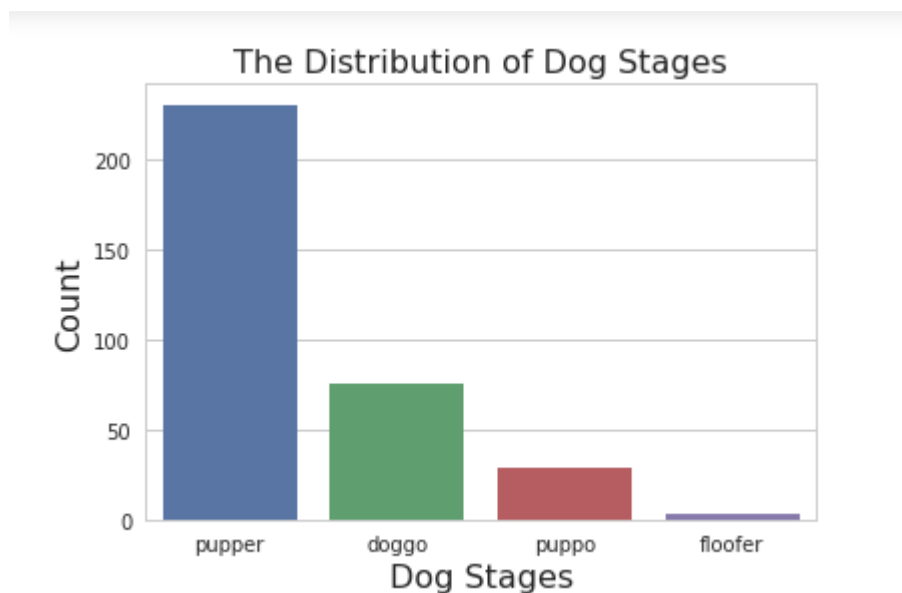
2- The Distribution of tweets sources.

The plot shows that the main source of tweets is from twitter app for iPhone while the TweetDeck is very rare [less than 1%].
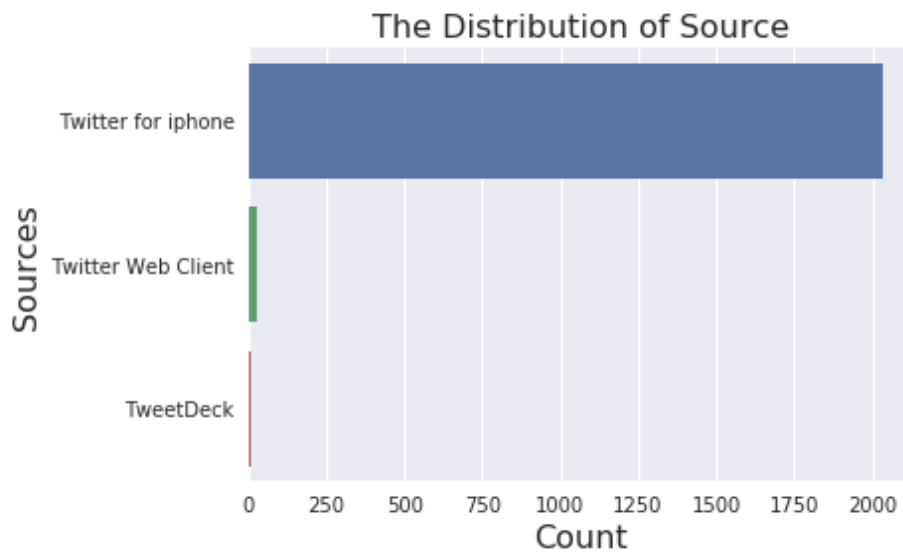
3- Relationship between favorite and retweet counts.

The plot shows a strong positive correlation between retweet counts and favorite counts, which means if a tweet has a high favorite count it will also have a high number of retweets.