Analysis & Visualization
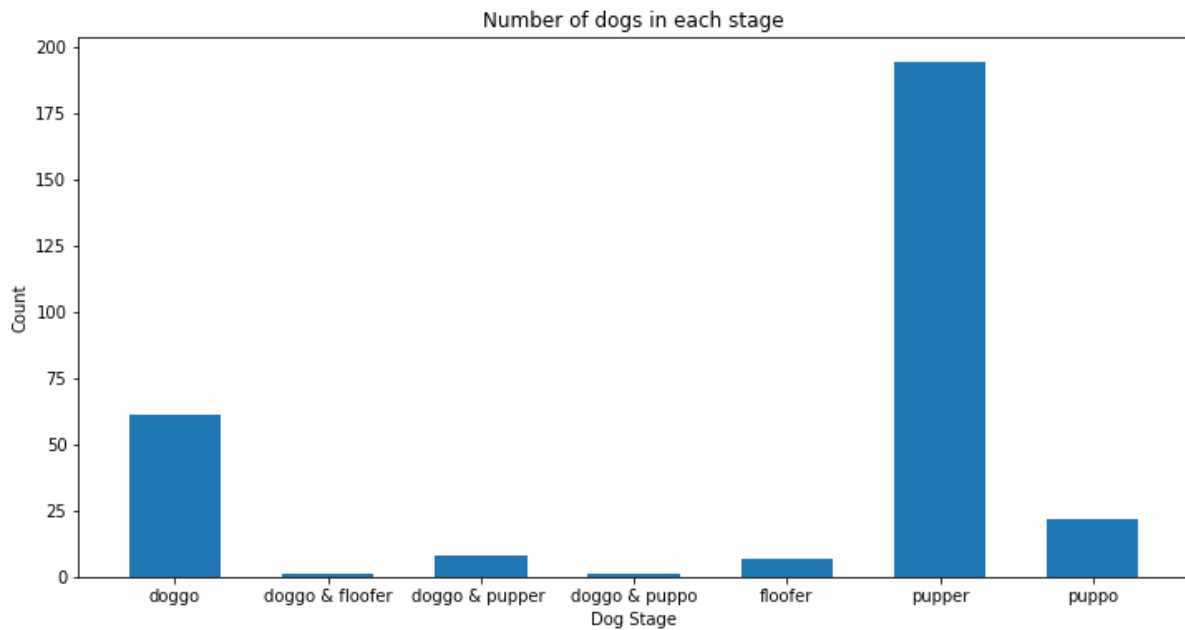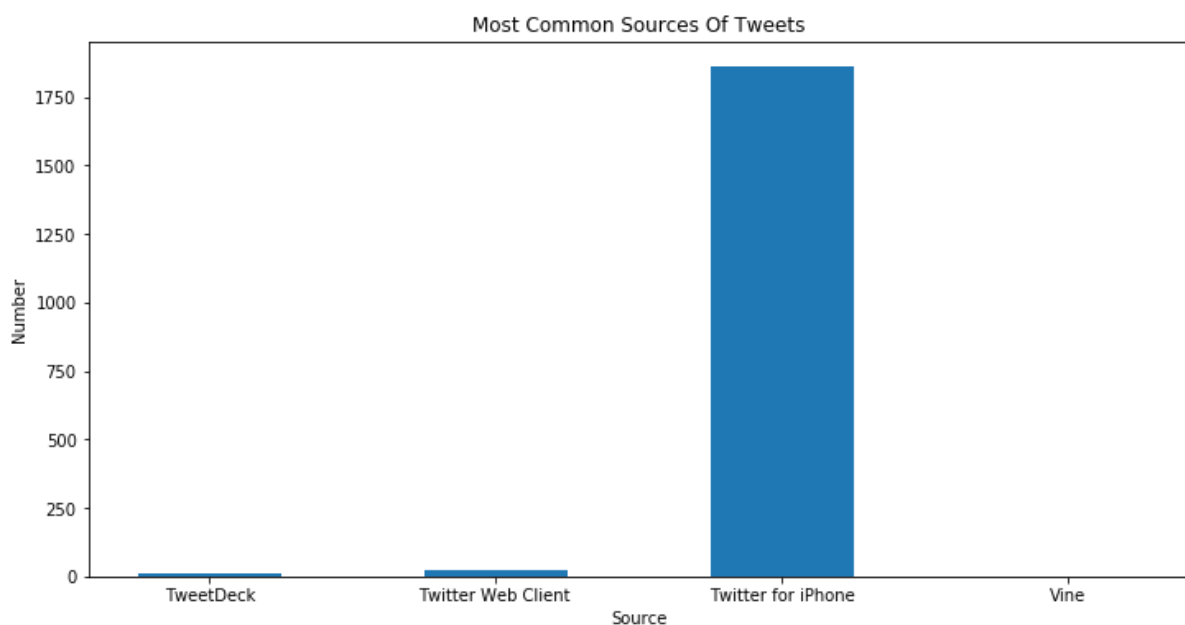
In this report I display 2 visualizations:

1- The Most Common Dogs Stage.

2- The Most Common Sources Of Tweets.

Insights

- The most common Dog Stage is 'Pupper' and the least is 'Doggo & Floofer' and 'Doggo & Puppo.

- The most common source of tweets is 'Twitter for iPhone' with a number of 1861 while the least is ' TweetDeck' with 10 times.

- The Mean Retweet Count is 2801.

- The Mean Favorite Count is 9043.

- The Minimum Image Number is (1) & The Maximum is (4).

- There Are some incorrect ratings I noticed, some of them were extracted incorrectly from the text, others were unfortunately extracted right but received low values such as '1/10', '2/10', '5/10'. It turned out they weren't dog images!
- I also noticed while visually checking the ratings that the majority lies within this range: {12-13-14} with 13 being the most common rating, surprisingly there was a 420/10 rating!! But when i saw the image of that rating i was no longer surprised:

Here's the image with 1/10 rating:

Here's another one with a 2/10 rating:

We can see from this analysis that the Twitter account 'We Rate Dogs' is quite popular worldwide and this made the analysis much easier since we have multiple parameters and features to measure and also criteria to compare against each other for eg. We created a plot relating each dog stage to the number of dogs existing in this stage from which we concluded that the 'Pupper' stage is the most prevailing one in our dataset while the ''Doggo & Floofer' and 'Doggo & Puppo are the least frequent.

In the second visualization we can see the most common source for tweets is 'Twitter for iPhone' and the least one is 'Tweet Deck' and this makes sense since a lot of people actually don't use Tweet Deck.

I've noticed that some tweets have very low ratings which is not common since we all love dogs.

We could've also made use of the other features in our dataset such as the ratings columns, we can use a bar chart or a line chart to see the distribution of ratings and see which photos had the highest ratings.

We could also plot the Tweet Vs. Timestamp to check the activity of this account through a certain period of time and see when exactly people tweeted the most.