

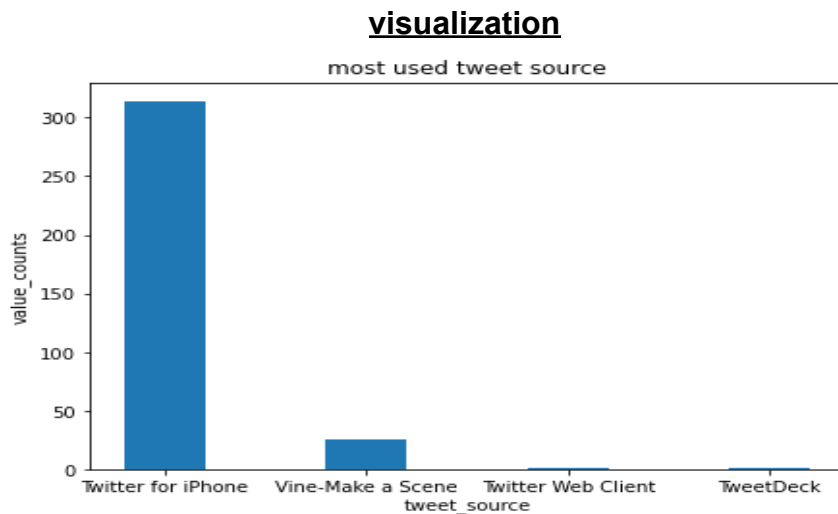
Investigate a dataset (weratedogs)

insights and visualization(s)

I made 4 insights to the dataframe:

1. Which tweet_source was the most used?

With the value count of 314, twitter for iphone was the most used source by a large gap. followed by vine-make a scene and the least being twitter web client and tweetdeck with the count of 2 each.

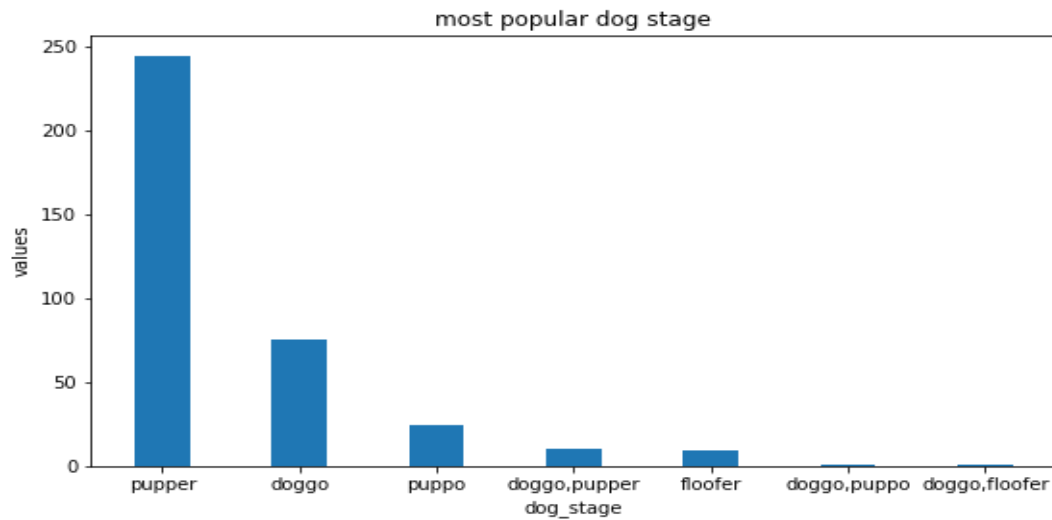


This insight was made to figure out which dog_breed was the most popular in the dataframe using matplotlib to create the barchart and plt.bar()

2. Which dog breed is most popular ?

The most popular dog stage is pupper with 244 counts and the least is 'doggo,puppo' and 'doggo,floofer' with count of 1 each.

Visualization

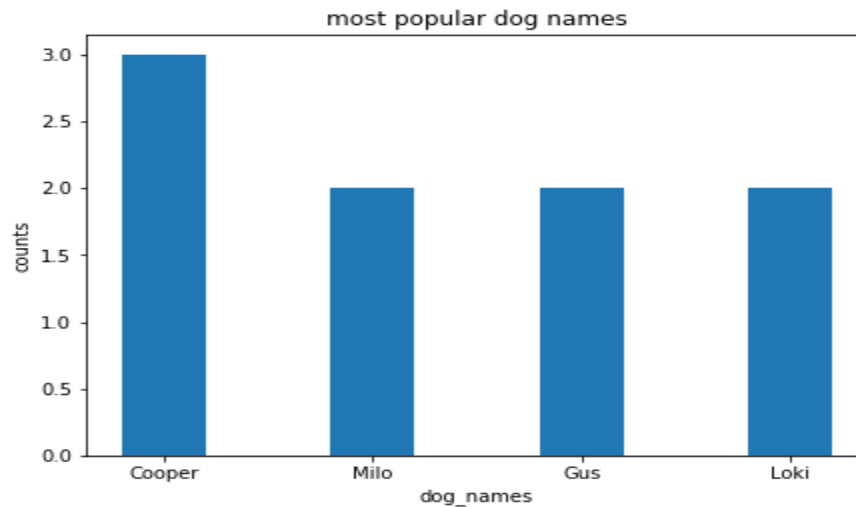


A picture of a pupper from reddit.

3. **Which is the most popular dog name?**

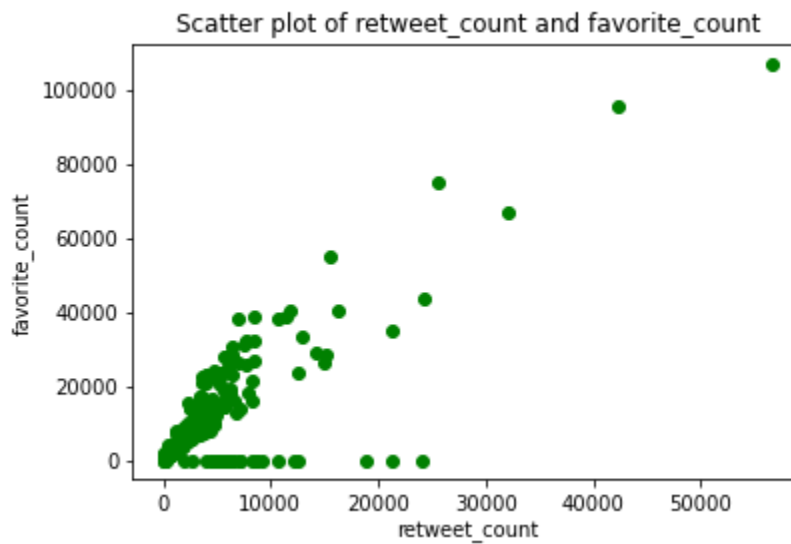
Cooper is the most popular dog name with a count of 3 and the least are Milo, Gus and Loki with counts of 2 each

visualization



4. Correlation between retweet_count and favorite_count

visualization



From the scatter plot, there is a positive correlation between retweet_count and favourite_count. This can be interpreted to mean that the more people like a tweet, the more retweets it gets. As one increases, so does the other.

Finally, let's see some pictures of the dogs posted in the weratedogs twitter page.



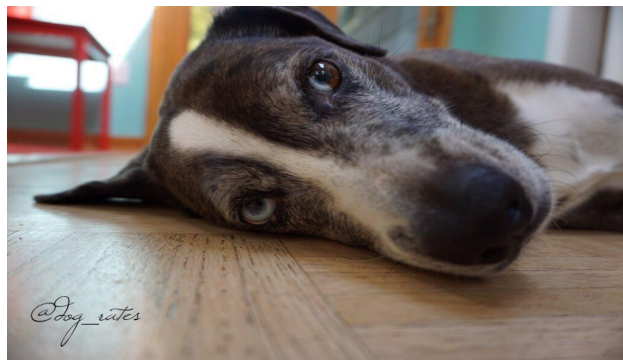
This cute dog is called Koda. He's a flat coated retriever.



This is Sandy(with the cute eyes and shirt) and she's a pembroke



This cute dog is named Pablo. He's a labrador retriever and a puppo



This cute dog is named Bayley. She's a badger and a puppo.