Analyzing the data:

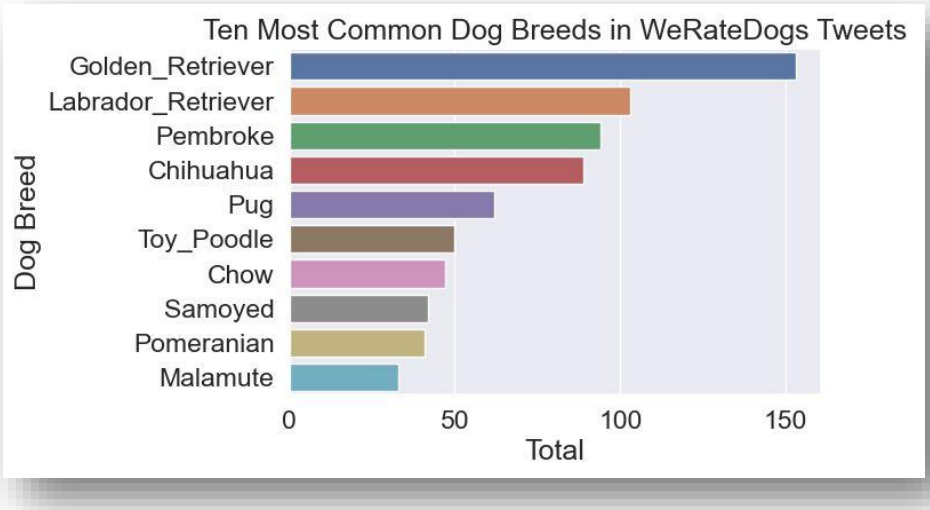
Analyzing these data set wasn't easy for me especially as the twitter account developer get accepted late which makes me working day and night to provide the project as early as possible.

So in the analyzing I start with the describe() method to help me get an overview of the data, from which I knew the average dog rating and the 5 value that describe any numerical columns as show in the next table:

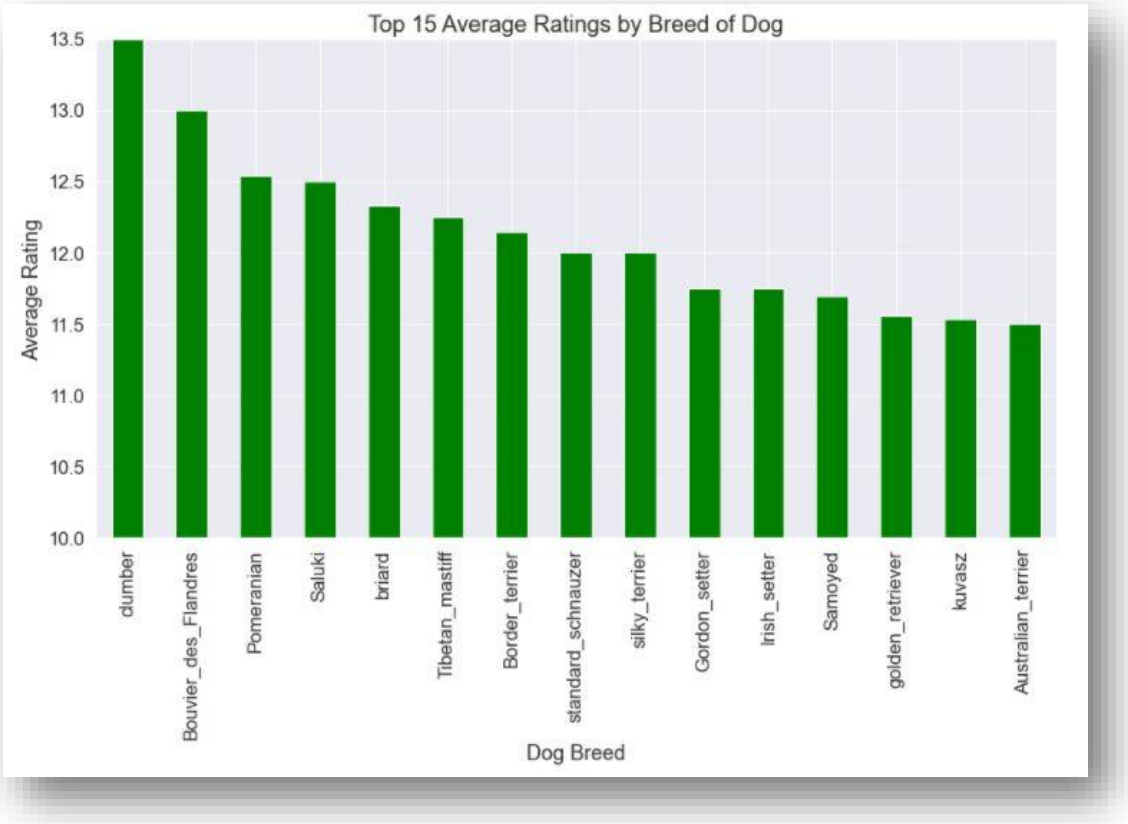
	tweet_id	dog_rating	rating_from_10	probability	no_of_likes	retweet_count
count	1.954000e+03	1954.000000	1954.0	1954.000000	1944.000000	1944.000000
mean	7.361899e+17	11.700614	10.0	0.465243	8243.746399	2435.923354
std	6.768652e+16	41.079593	0.0	0.339579	12128.817284	4355.572812
min	6.660209e+17	0.000000	10.0	0.000000	70.000000	11.000000
25%	6.758214e+17	10.000000	10.0	0.140000	1744.250000	541.000000
50%	7.087745e+17	11.000000	10.0	0.460500	3722.000000	1177.500000
75%	7.883537e+17	12.000000	10.0	0.775750	10325.250000	2789.750000
max	8.924206e+17	1776.000000	10.0	1.000000	154385.000000	76622.000000

After that I get **Ten Most Common Dog Breeds in WeRateDogs Tweets**

Which I found that the Golden Retriever, Labrador Retriever and Pembroke get in the top of the list as shown in the next figure:

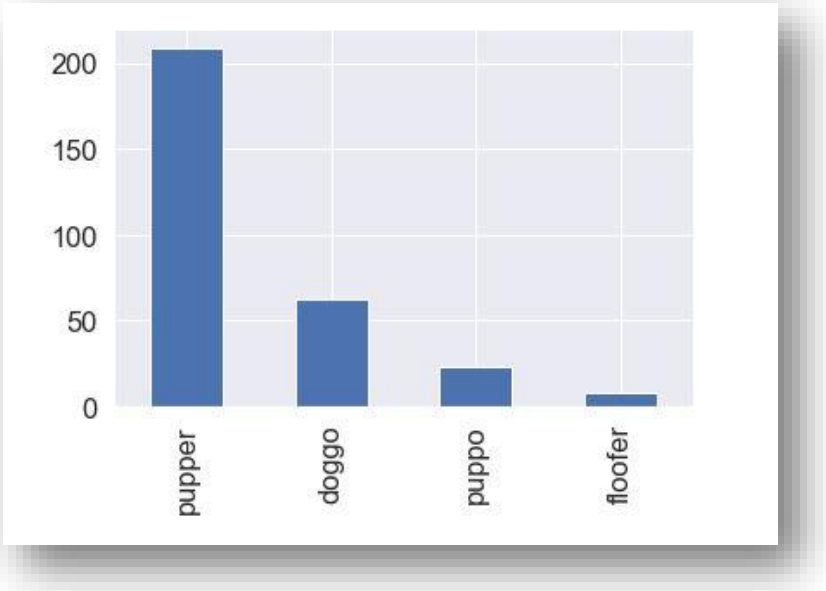


After that I want to know **what the top average rating by bread is** so I grouped the dog bread and then plot its data and found the dumber, Bouvier des Flasnders and Pomeranian in top of this list.

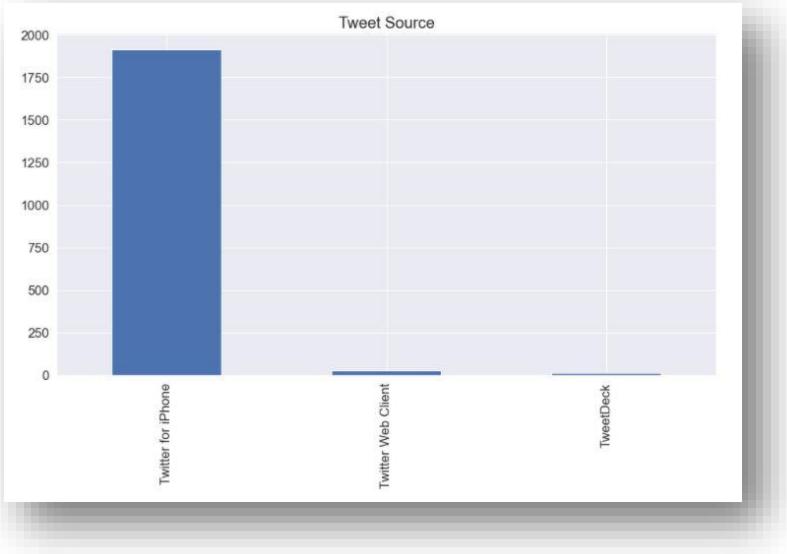


And I found **the number of every dog stage** excluding the none ones.

And plot it as bar and pie and get the next figures:



I found the **count for every tweet source** for the tweet to know the target if they want to expand and want to know which device they will develop for:



I think that there's a lot information can be extracted from these data set but the time is too limited as I don't want to be late more than that.