

Mainly my analysis target the relation of different variables with ratings and likes given to these dogs through the following questions and then I'll explain the technical ways and the small tasks I'll use to answer these questions.

Questions:

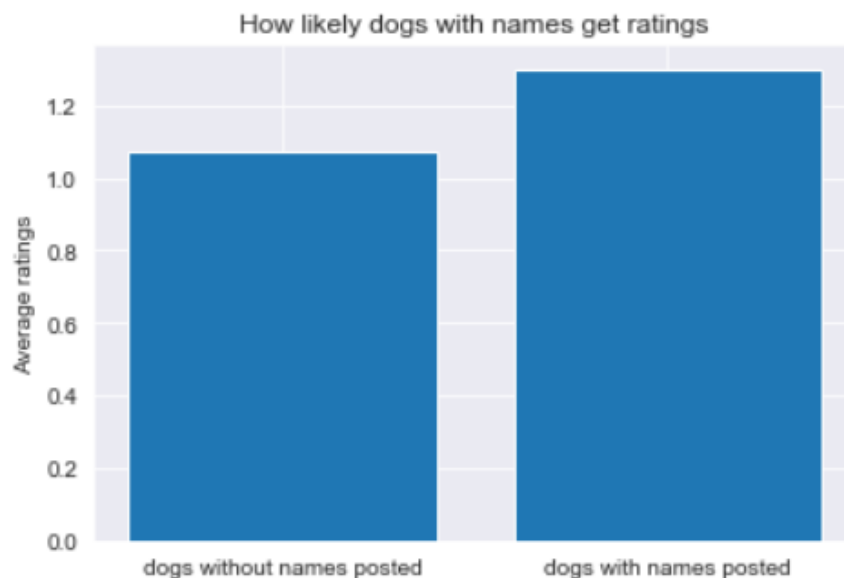
- 1- Is there a relation between the dog's name mentioned and how likely that will get it a high rating?
- 2- What breeds get more high ratings?
- 3- What breeds get more likes and retweets?
- 4- Do the tweet time make a difference in likes and retweets?
- 5- The dog physical and age features "the dog dictionary", which features get the highest ratings?
- 6- What's special about the dogs and tweets having these rare ratings out of range like 600/10?

Analysis and visualizations:

- 1- Statistics says that the most owned breeds are:
 - a) Golden retriever
 - b) Labrador retriever

And that golden retrievers get high rates.

- 2- Let's start and see if sending the dog's name or mentioning it may affect the rating it's going to get



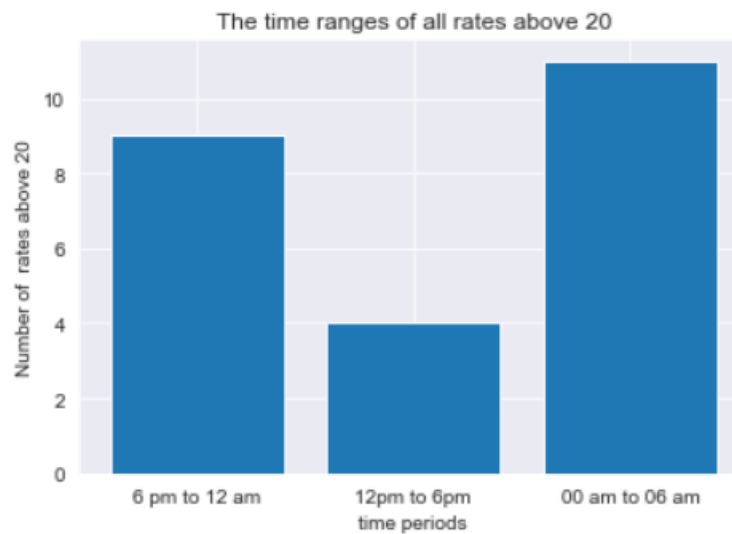
As we can definitely see, when the dog's name is mentioned it's more likely to get higher rating 17.6%

This is some closer look at the data used

:

	tweet_id	in_reply_to_status_id	in_reply_to_user_id	retweeted_status_id	retweeted_status_user_id	rating_numerator	rating_denominator	rating_value
count	1.601000e+03	1.000000e+00	1.000000e+00	1.150000e+02	1.150000e+02	1601.000000	1601.0	1601.000000
mean	7.466872e+17	6.678065e+17	4.196984e+09	7.756884e+17	4.160877e+09	10.739538	10.0	1.073954
std	6.877601e+16	NaN	NaN	5.923931e+16	3.872055e+08	1.937522	0.0	0.193752
min	6.660293e+17	6.678065e+17	4.196984e+09	6.671522e+17	4.466750e+07	2.000000	10.0	0.200000
25%	6.801611e+17	6.678065e+17	4.196984e+09	7.408721e+17	4.196984e+09	10.000000	10.0	1.000000
50%	7.327322e+17	6.678065e+17	4.196984e+09	7.833346e+17	4.196984e+09	11.000000	10.0	1.100000
75%	8.047388e+17	6.678065e+17	4.196984e+09	8.191170e+17	4.196984e+09	12.000000	10.0	1.200000
max	8.924206e+17	6.678065e+17	4.196984e+09	8.874740e+17	4.196984e+09	14.000000	10.0	1.400000

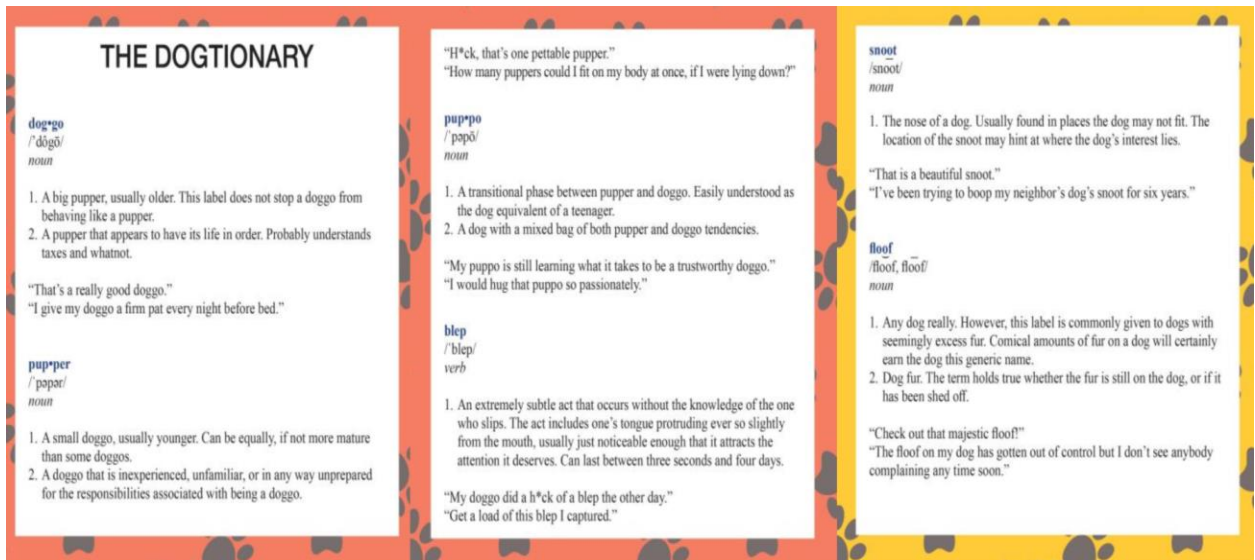
3- Let's have a closer look at these 24 values that got very high ratings "above 20"



Lovely notice the most of the very highest ratings was made and posted after midnight as the past bar chart shows three periods of the day with a small notice, no such high ratings has been posted between 6 am to 12 pm.

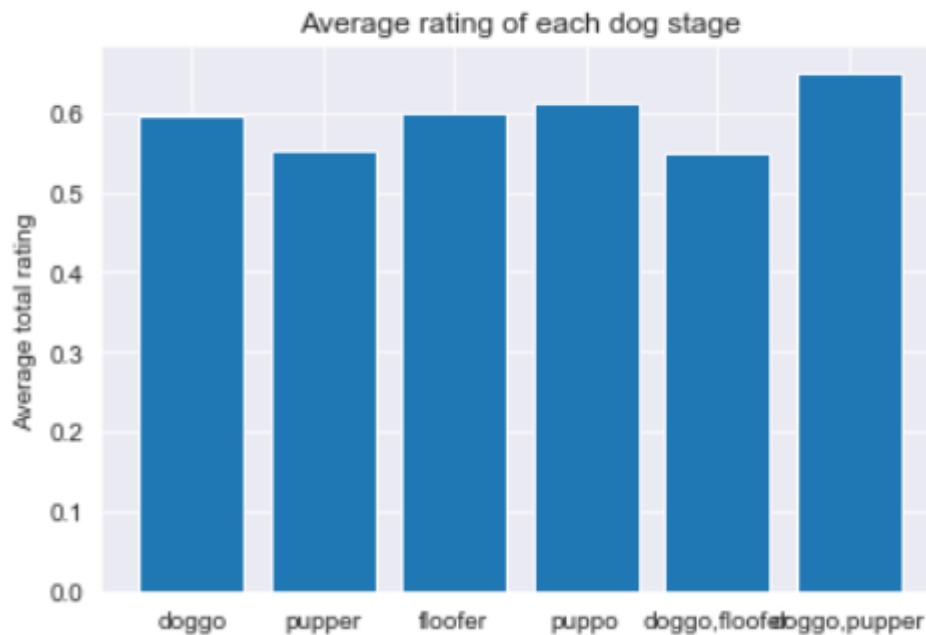
4- Let's see the statistics related to the dogs dictionary

Let me just show you the dictionary first:



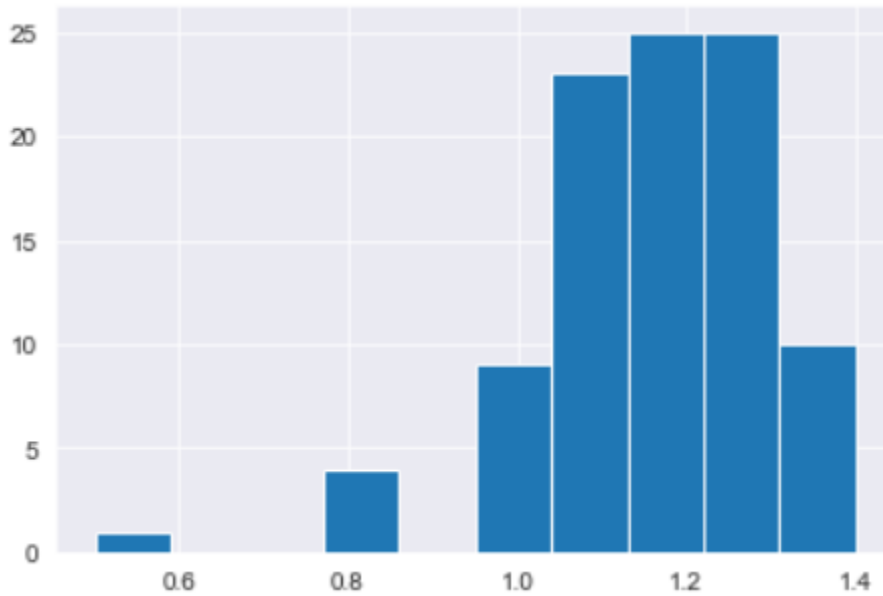
Well , which dog get the highest ratings in that dictionary ? Let's see

Text(0, 0.5, 'Average total rating')



We got a winner , Dogs in the doggo,pupper stage always get higher ratings by a slight difference after the puppo as an individual stage.

Closer look at one of the histograms used in the previous analysis testing



It shows variation of the ratings of the doggo dog type.