



## DOGGOS, FLOOFERS, PUPPERS AND PUPPOS

A look through 'We rate dogs' twitter account.

Mohammad Tawfik

Data analysis – professional track

Looking through the data and with some basic analysis, very interesting information were found regarding 'We rate dogs' twitter account.

## Insights

Looking at this table:

	rating_numerator	rating_denominator	retweet_count	favorite_count
dog_stage				
doggo	11.830986	10.000000	5905.802817	17220.873239
doggo-floofer	11.000000	10.000000	2991.000000	15600.000000
doggo-pupper	11.111111	10.000000	3795.666667	12355.666667
doggo-puppo	13.000000	10.000000	17041.000000	43780.000000
floofer	11.888889	10.000000	3562.222222	9913.000000
pupper	10.809091	10.000000	2151.681818	6594.804545
puppo	12.000000	10.000000	5796.000000	20565.000000
NaN	12.389047	10.536794	2304.650314	7828.237307

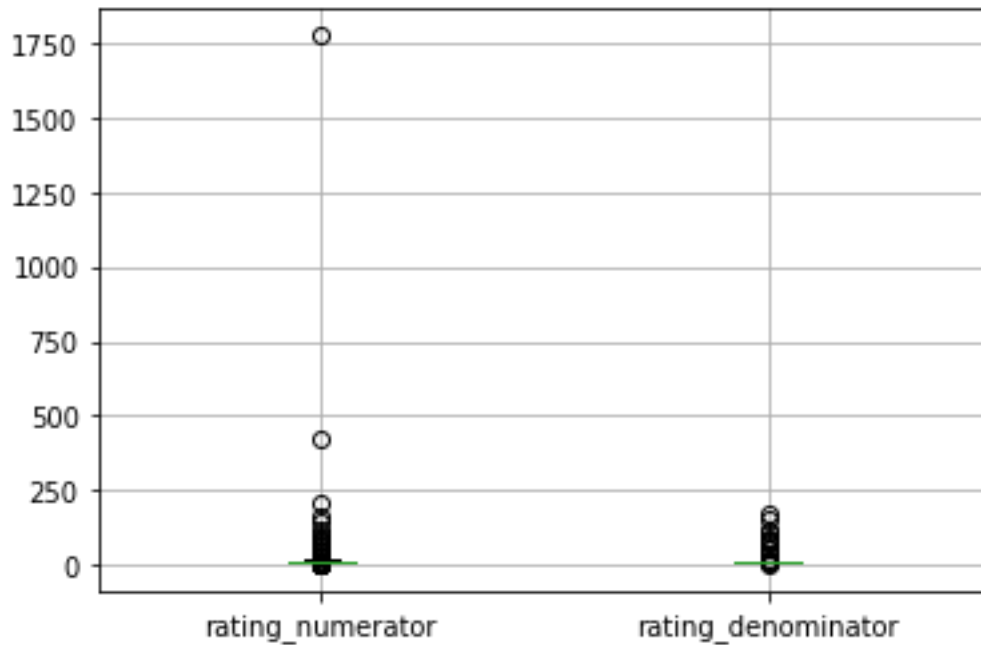
We can see that tweets that had more than one dog with dog stage 'doggo-puppo' were most popular with people and had the most both retweet counts and favorite counts. On the other hand, tweets with dog stage 'pupper' had the least retweet counts and favorite counts, even less than tweets with no dog stage at all although they are very close. 'doggo-puppo' also had the highest average rate.

While this table:

		retweet_count	favorite_count
year	month		
2017	8	6543.500000	33153.500000
	7	6175.274510	29876.666667
	6	5655.659574	26224.702128
	5	6768.534884	27322.116279
	4	5026.095238	21175.571429
	3	4134.200000	18210.680000
	2	3652.924242	16356.242424
	1	5341.536232	17407.724638
2016	12	4985.929825	15251.508772
	11	4131.500000	13076.814815
	10	4073.788732	12622.042254
	9	3087.455882	9891.411765
	8	3110.098361	9194.147541
	7	2469.610526	6817.442105
	6	3818.644444	9480.155556
	5	1973.355932	5767.593220
	4	1405.066667	4256.666667
	3	1523.878788	4196.924242
	2	1357.775000	3792.258333
	1	1444.961749	3797.426230
2015	12	1334.129380	3241.846361
	11	477.168919	1158.324324

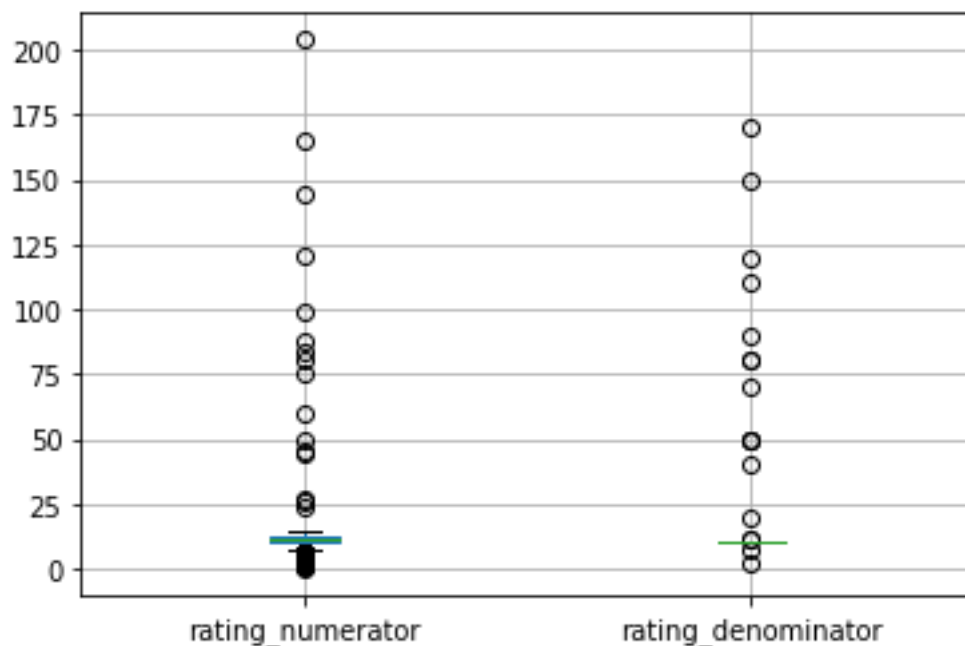
Shows that with through time, the popularity and engagement with the account increased very rapidly.

Looking at the box plot of rating numerator and denominator:



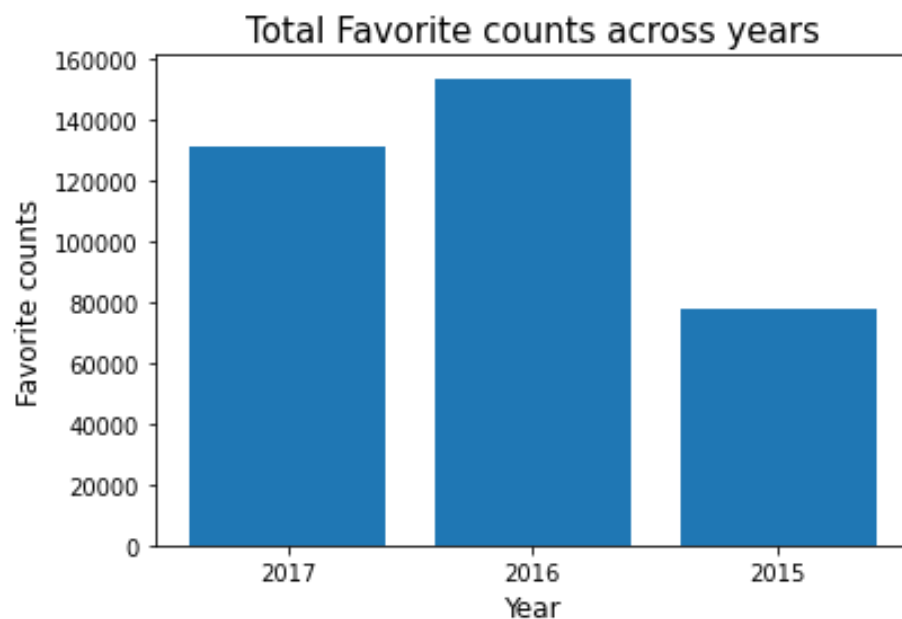
We can see 2 outliers in rating numerator box plot, but by checking them out it turns out that they are not really outliers, instead, they are just an expressive way to demonstrate how cute the dog is.

But let's remove them anyway for a second to have a better look at the distribution:



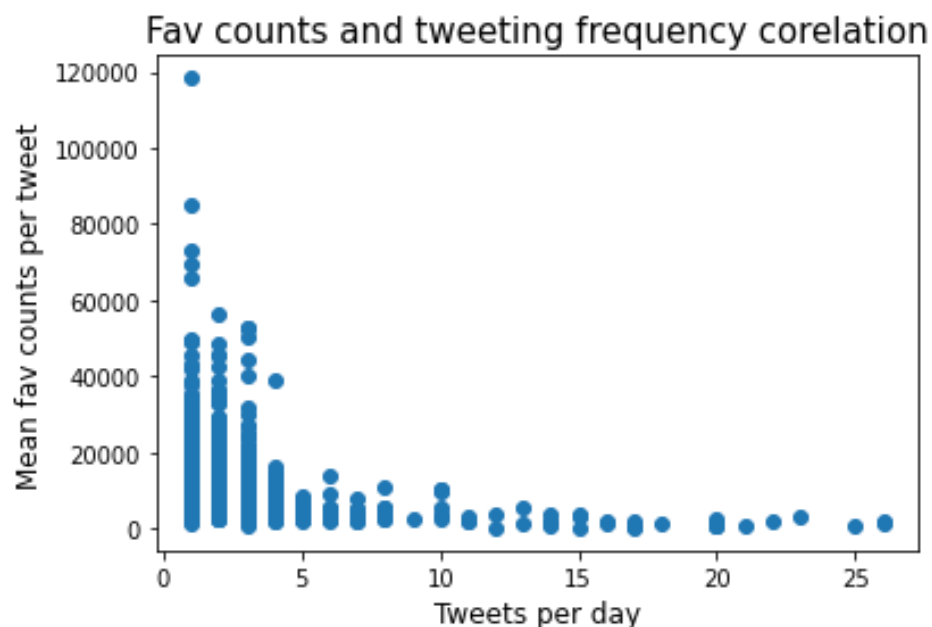
## Visualization

Looking at the next bar plot:



We can see that year 2016 had the highest favorite counts and 2017 was a little lower.

This scatter plot:



Tells us that also tweeting frequency per day decreased with time, the engagement didn't decrease, on the contrary it did increased, which indicates that people are really interested in this account.