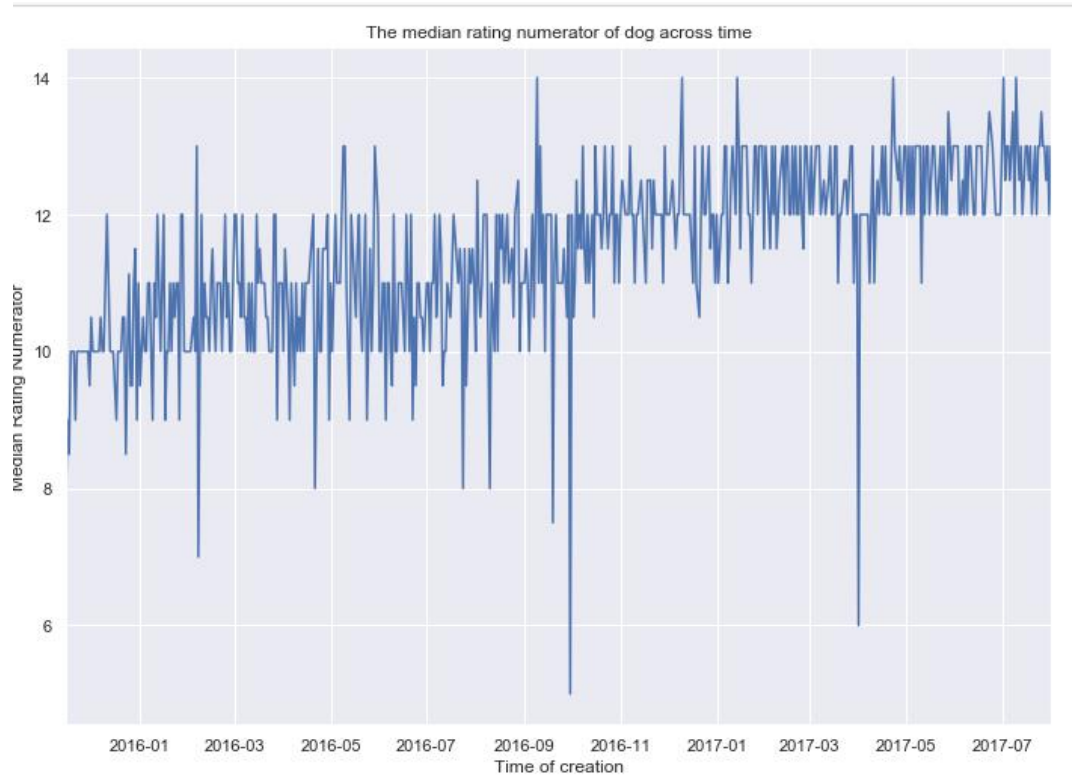


Act report.

The purpose of this document is to report some insights and vizualizations about WeRateDogs Twitter account we found.

We saw that the rating numerator had a median rating numerator that in general lines has grown over time. The median rating numerator in 2016-01 was between 9 an 12, in 2017-07 was between 12 and 14!

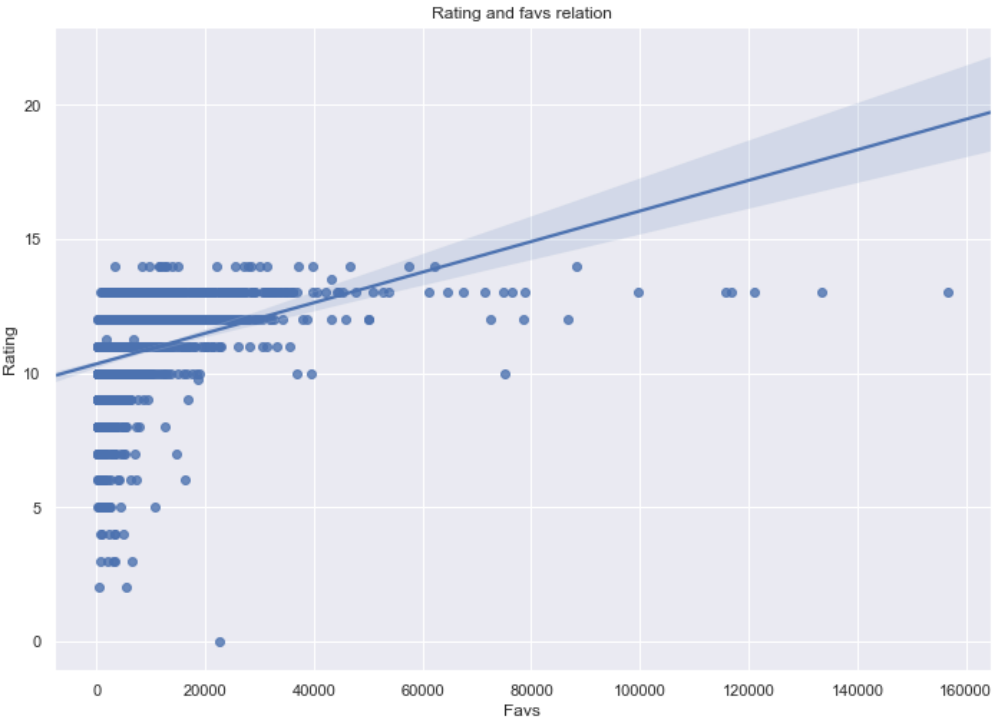
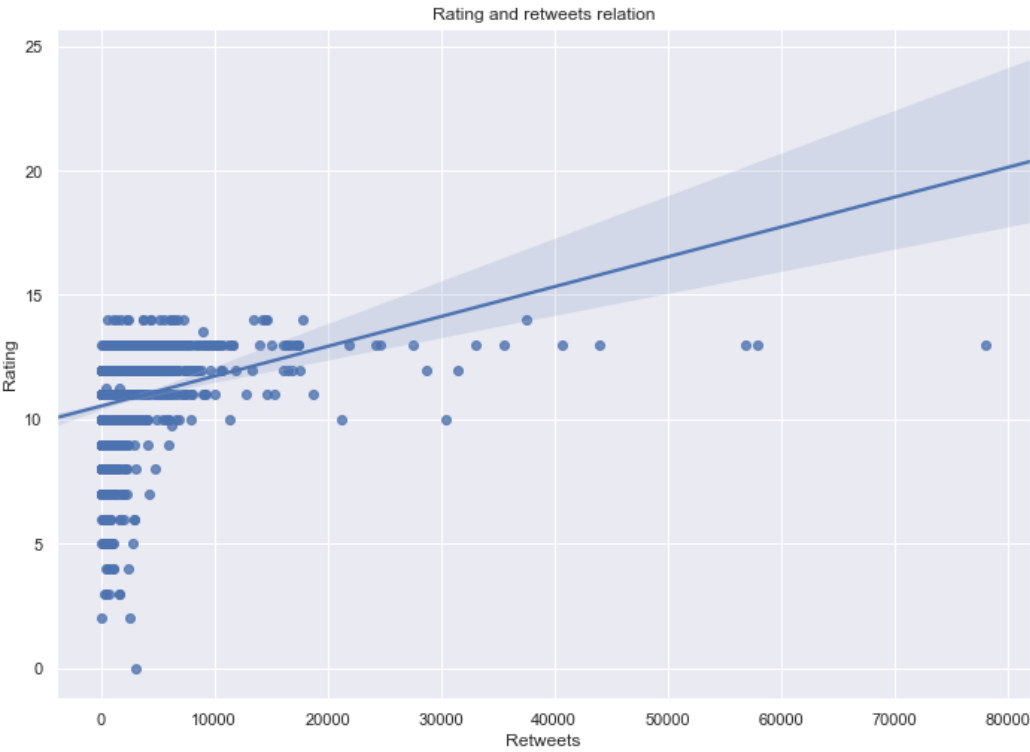


Then we tried to see if there were linear relations between variables.

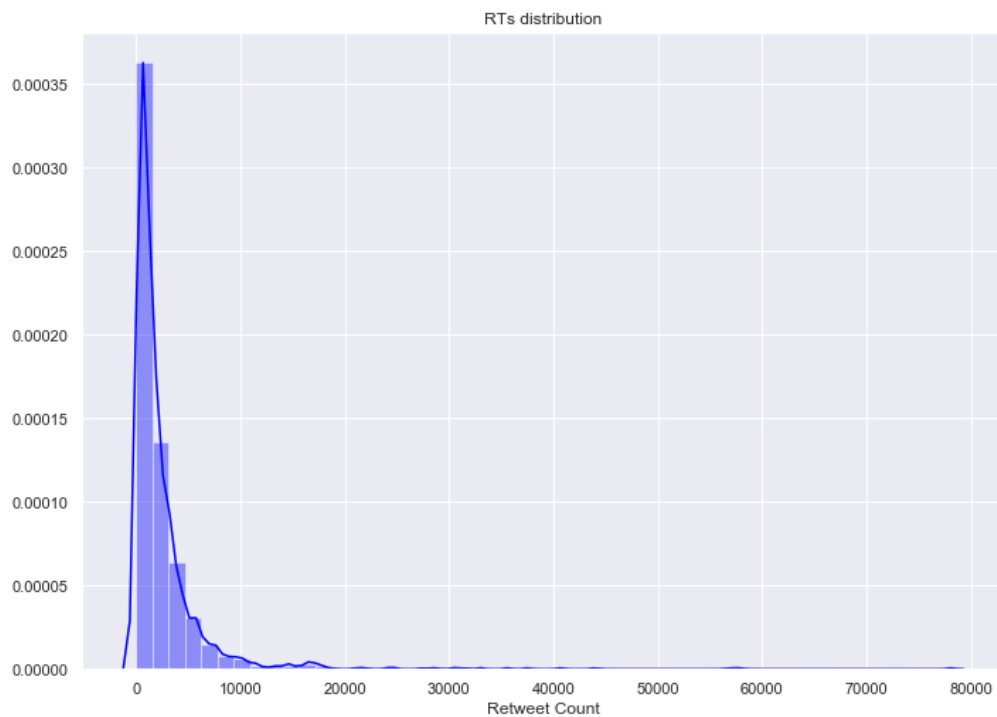
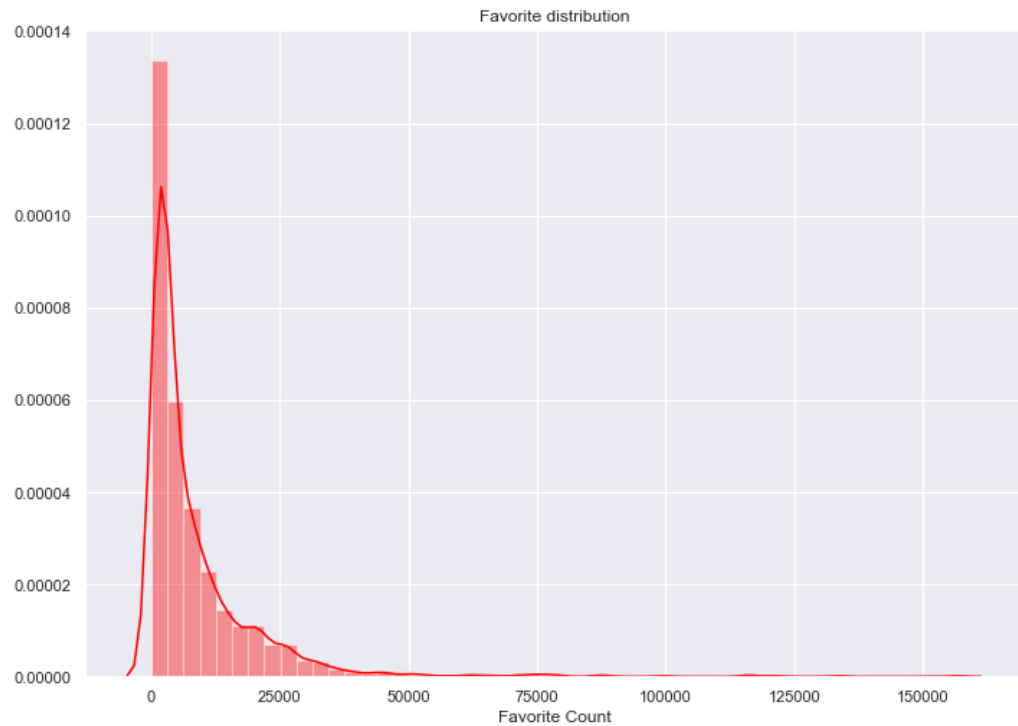
First we though about Rating_numerator and Retweets and Favs. We found a very small corr between rating and retweets and ratings vs favs). But then we found a huge relation correlation between retweets and favorites (93%!).

	tweet_id	rating_numerator	rating_denominator	p1_conf	p1_dog	p2_conf	p2_dog	p3_conf	p3_dog	retweets	favorites
tweet_id	1.000000	0.544031	NaN	0.112550	0.052143	-0.004696	0.056746	-0.055592	0.020532	0.374652	0.601078
rating_numerator	0.544031	1.000000	NaN	0.108134	0.133518	0.001538	0.045025	-0.044856	0.033827	0.307505	0.406601
rating_denominator	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
p1_conf	0.112550	0.108134	NaN	1.000000	0.089728	-0.539225	0.117166	-0.732677	0.079258	0.053042	0.075434
p1_dog	0.052143	0.133518	NaN	0.089728	1.000000	0.100744	0.213736	0.016627	0.111019	-0.016555	0.004582
p2_conf	-0.004696	0.001538	NaN	-0.539225	0.100744	1.000000	0.079687	0.477851	0.031204	-0.015214	-0.017083
p2_dog	0.056746	0.045025	NaN	0.117166	0.213736	0.079687	1.000000	-0.007387	0.085168	-0.007761	0.012149
p3_conf	-0.055592	-0.044856	NaN	-0.732677	0.016627	0.477851	-0.007387	1.000000	0.000107	-0.041999	-0.050904
p3_dog	0.020532	0.033827	NaN	0.079258	0.111019	0.031204	0.085168	0.000107	1.000000	-0.009502	-0.005875
retweets	0.374652	0.307505	NaN	0.053042	-0.016555	-0.015214	-0.007761	-0.041999	-0.009502	1.000000	0.930613
favorites	0.601078	0.406601	NaN	0.075434	0.004582	-0.017083	0.012149	-0.050904	-0.005875	0.930613	1.000000

Next two scatterplots present evidence of a weak but positive relationship between rating and retweets and rating and favs.

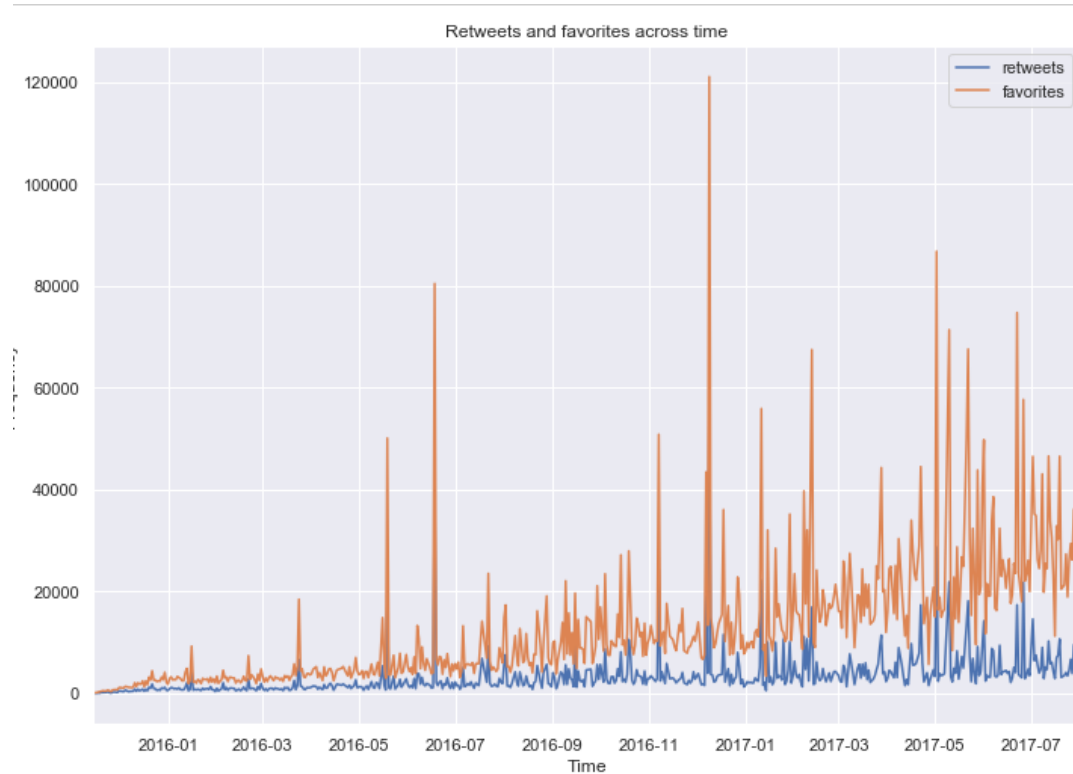


After confirming that we investigated a little more about Favs and Retweets. See the distplots below.



We know values were different (Favorites have a median near 4k and RTs near 1.2k), but they had a similar behavior pattern.

Then we figured if they had the same behavior across time. Seems yes but with favs have a more pronounced behaviour



Finally we did a Scatterplot with a regline showing a huge correlation between variables.



As a bonus track we investigated about prediction confidence. The first predictor had a huge confidence much bigger than the second and the third. Even has predictions near 100% of confidence.

