'We Rate Dogs' tweets Analysis

Noura Shebl

It was both fun and interesting to analyse the tweets from WeRateDogs tweets. This is a Twitter account that rates people's dogs with a humorous comment about the dog. It has a unique rating system, since it rates dogs with values in the numerator greater than the denominator, such as 12/10. Another interesting fact is that this all started only since 2015, yet it achieved a world wide impact.


At the beginning I wanted to have a general idea using statistical summaries such as mean, and median.

|	rating_numerator	rating_denominator	favorite_count	retweet_count	img_num	p1_conf	p2_conf	p3_conf	ratio
count	1994.000000	1994.000000	1994.000000	1994.000000	1994.000000	1994.000000	1.994000e+03	1.994000e+03	1994.000000
mean	12.280843	10.532096	8895.725677	2766.753260	1.203109	0.593941	1.344195e-01	6.024848e-02	1.169249
std	41.497718	7.320710	12213.193181	4674.698447	0.560777	0.271954	1.006807e-01	5.089067e-02	4.067071
min	0.000000	2.000000	81.000000	16.000000	1.000000	0.044333	1.011300e-08	1.740170e-10	0.000000
25%	10.000000	10.000000	1982.000000	624.750000	1.000000	0.362857	5.393988e-02	1.619283e-02	1.000000
50%	11.000000	10.000000	4136.000000	1359.500000	1.000000	0.587635	1.174550e-01	4.950530e-02	1.100000
75%	12.000000	10.000000	11308.000000	3220.000000	1.000000	0.846285	1.951377e-01	9.159438e-02	1.200000
max	1776.000000	170.000000	132810.000000	79515.000000	4.000000	1.000000	4.880140e-01	2.734190e-01	177.600000

From this brief analysis, I noticed that the mean for both favourite counts and retweet counts was much greater than the median. The mean is almost double the median. This strongly indicated that the data is skewed to the right.

I then plotted the changes in favourite counts against the time frame. This time covered a duration span of about two years from 2015 to 2017. And as

expected, the favourite count had an increasing trend over time. This is because with time twitter users increase as well as dog lovers. They are such loyal pets.
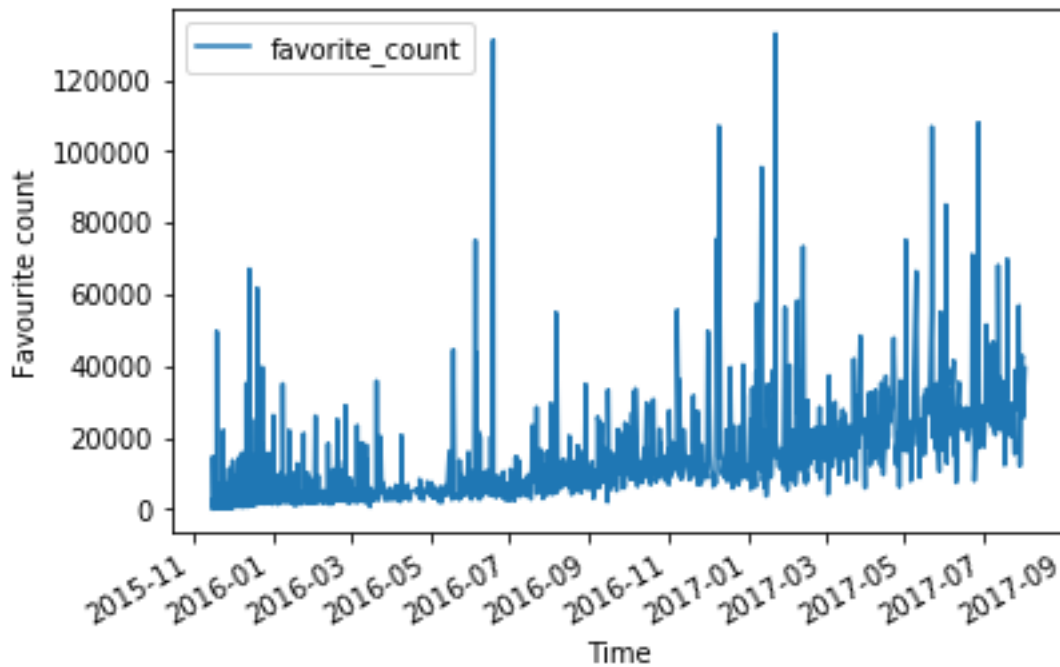
And finally, I wanted to explore the relationship between the favourite counts and the retweet counts.

Log(Retweet Count) versus Log(Favorite Count)
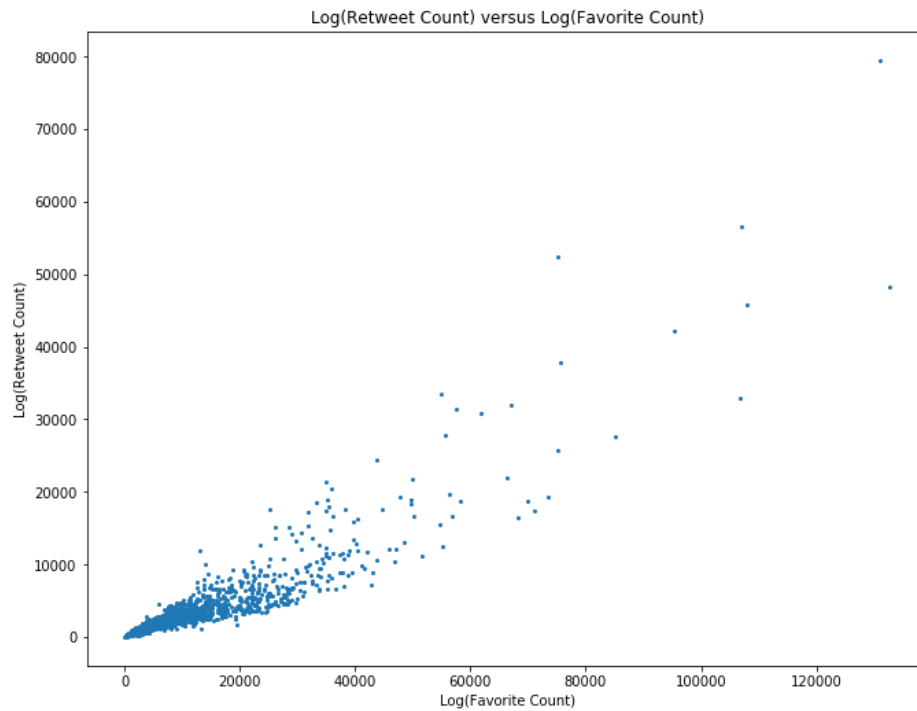
This graph showed a very strong positive correlation between both favourite counts and retweet counts. It is beautiful to visualize that using scatter plots. Yet it is not surprising, because among users of twitter, those tweets that are marked as favourite by users are mostly also retweeted, thus explaining this strong relationship between then two factors.