

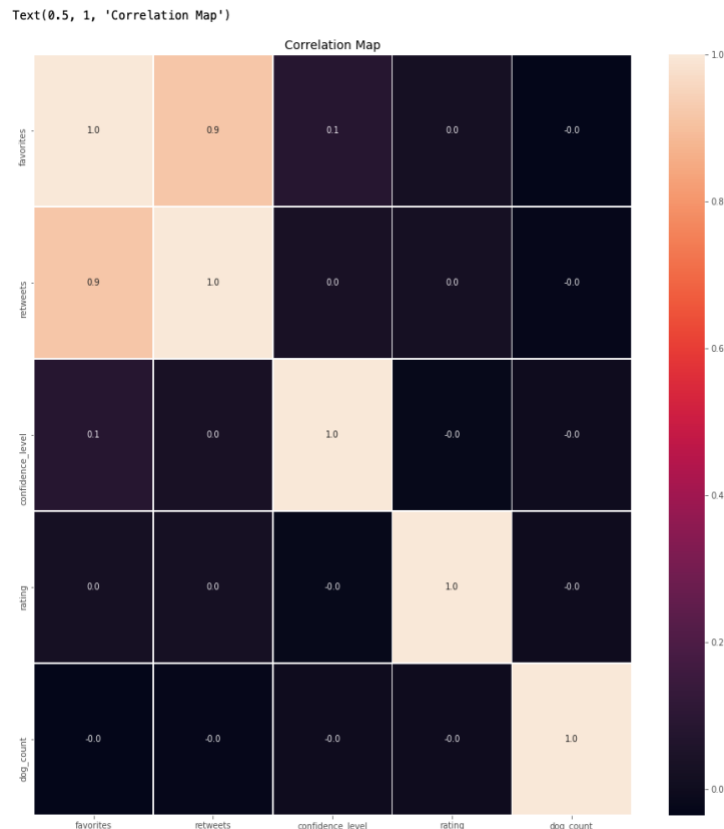
## Analyzing and Visualizing WeRateDogs

WeRateDogs is a Twitter account that supplies the platform to let people rate people's dog with a humorous comment about the dog. The ratings almost always have a denominator of 10. However, the numerators are almost greater than 10 (11/10, 12/10, 13/10, etc.)

Because WeRateDogs has over 6 millions followers, so I would like to share some interesting topics I found from there such as which dog type is the most common dog and what dog type has the highest average rating.

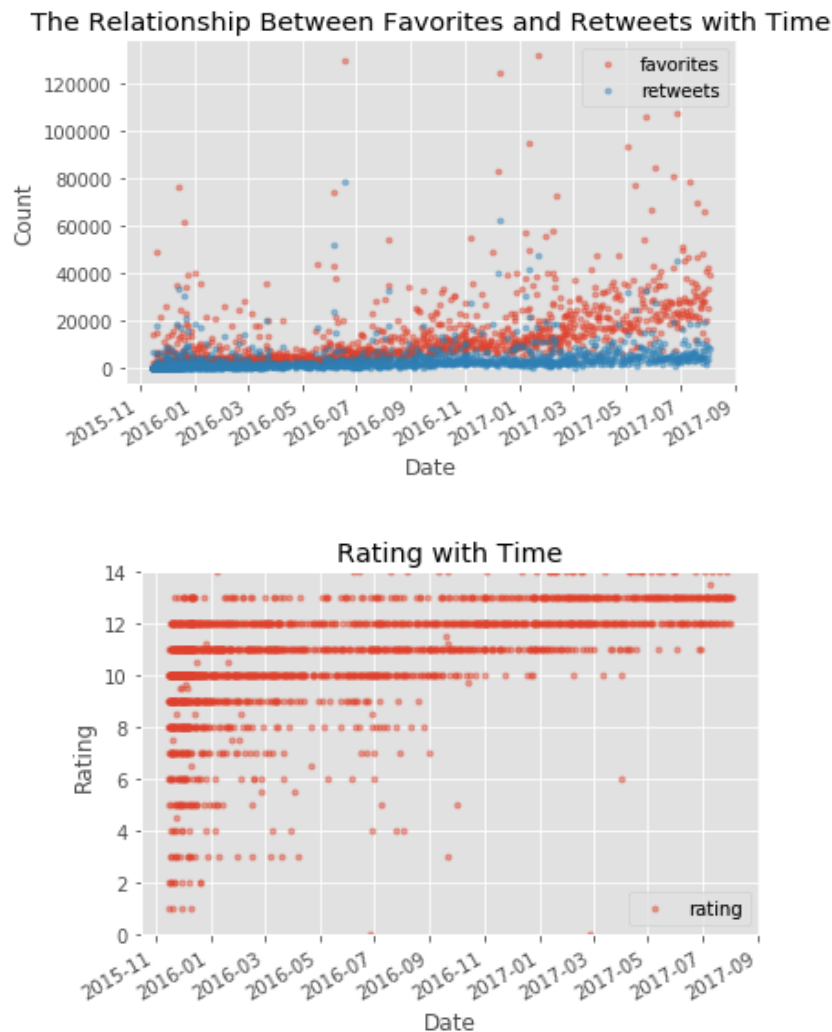


First of all, I use correlation map to see the relationships between variables as below,



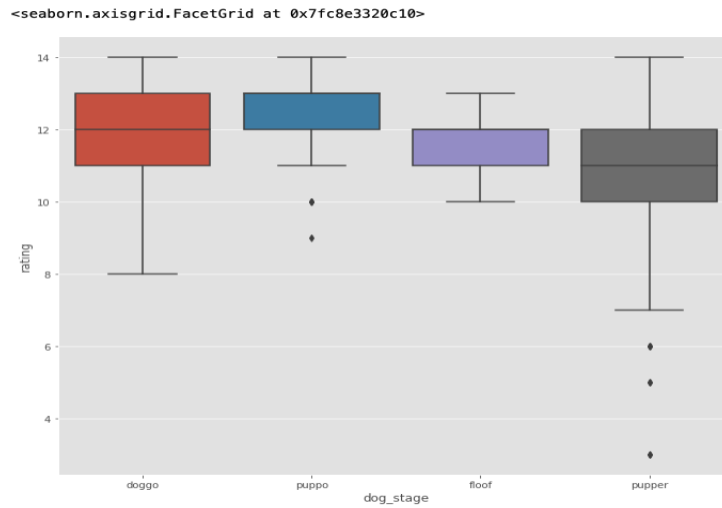
From above correlation map we can see that only the relationship between favorited and retweet is strong positive. The relationship between favorites and confidence\_level is also positive but weak.

Base on the map, I would like to do further analyses for favorites and retweets and rating with time as below,



According to above analyses we can see that retweets and favorites have strong positive relationship. In other words, if the dog will get more favorites and retweets with higher rating if the dog is getting better. However, we didn't see the relationship between rating and retweets or rating and favorites is strong and positive.

And then I would like to compare the relationship between rating and different dog stages. I used box plot to show the result as below,



According the above analysis, we found that pupper has lower rating and puppo has higher rating. And puppo has higher rating. And then I try to use describe() to get the further insight as below,

	count	mean	std	min	25%	50%	75%	max
<b>dog_stage</b>								
<b>None</b>	1624.0	11.766882	45.025499	0.0	10.0	11.0	12.0	1776.0
<b>doggo</b>	69.0	11.797101	1.510548	8.0	11.0	12.0	13.0	14.0
<b>floof</b>	34.0	11.705882	0.759961	10.0	11.0	12.0	12.0	13.0
<b>pupper</b>	237.0	10.616160	1.833623	3.0	10.0	11.0	12.0	14.0
<b>puppo</b>	29.0	12.172414	1.197288	9.0	12.0	13.0	13.0	14.0

In the previous analysis, we found that pupper has lower rating and puppo has higher rating. So, I tried to get more insight from above table. And we can see that although pupper has lower mean for rating, but the count for pupper is highest except none. The count of puppo is lowest although it has the highest mean for rating. For the rest of the dog stage, we do see that the counts are also significantly lower than pupper. So, I think it will be fair to compare once we have more counts for doggo, floof and puppo.