

# Analyzing and Visualizing Data

## Introduction

WeRateDogs Twitter account is a famous account that tweets about dogs and rates their images.

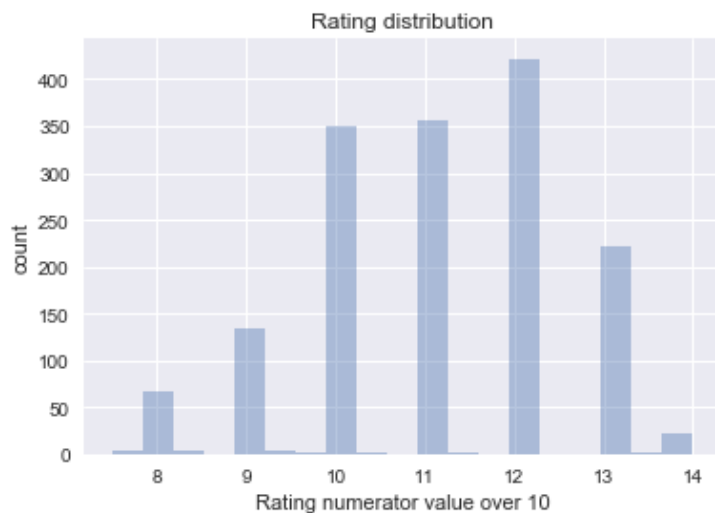
This account adopts a rating methodology where numerator can exceeds the denominator in most of the ratings and this what make this account so special.

In the Analysis, we look in more details about the rating distribution and whether the rating given by this account has any relationship with the number of favorite counts and retweet counts that each rated photo has gained overtime.

Then we will look deeper at the breed of the dogs that appear the most in the WeRateDogs Tweets.

## Analysis

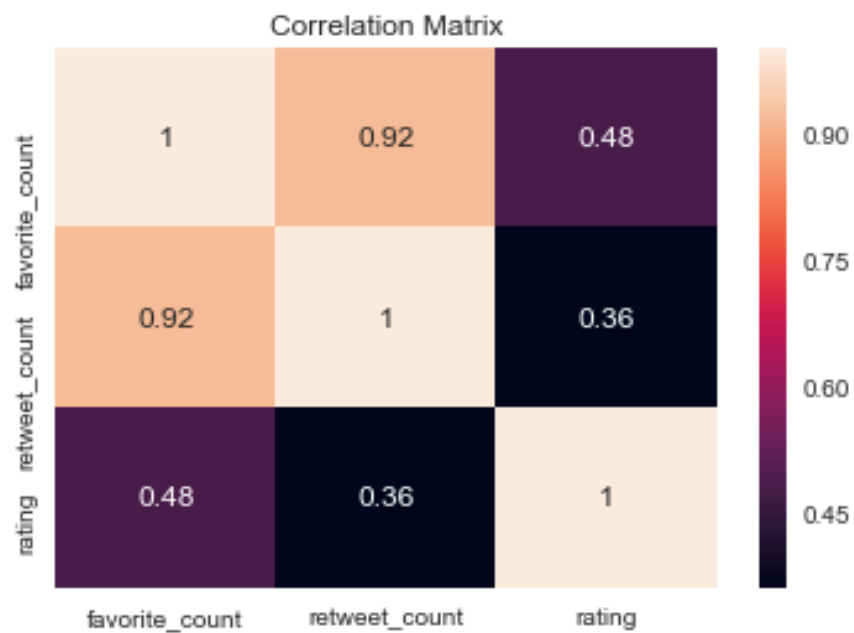
From the gathered info, if we take out outliers, we can see that most of rating falls between 10/10 to 14/10 with the mode is for 12/10.



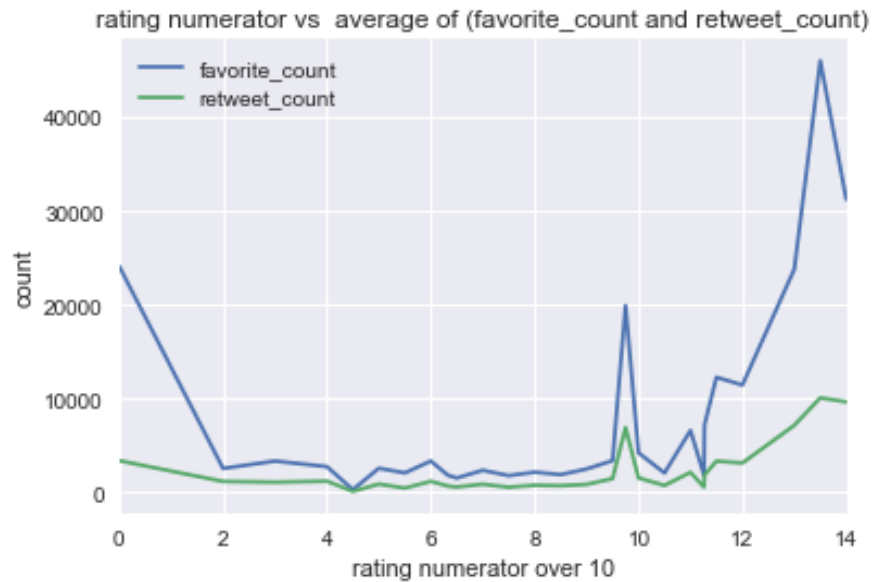
When looking at the rating what can be directly spotted is the huge rating of 1776/10! for a dog called Atticus celebrating America day, this would explain that the rating sometimes depends on the situation, the subject of the photo and not only the dog cuteness:



Moving on, from the correlation matrix, below we see a strong relationship between favorite counts and retweet count, but also we can notice a moderate correlation between favorite count and the rating.



Well let's dig deeper into this and plot rating numerator vs the Average of favorite count and retweet count:



Well when the rating goes above 11/10 to 13/10 we notice that the favorite count trend increases as well as the retweet count trend. But what's more interesting to look at is the spike below 10/10 in favorite count and retweet count which is a bit surprising. Well that's the tweet that gained a lot of traction though the rating is actually 9.75/10.



WeRateDogs™  
@dog\_rates

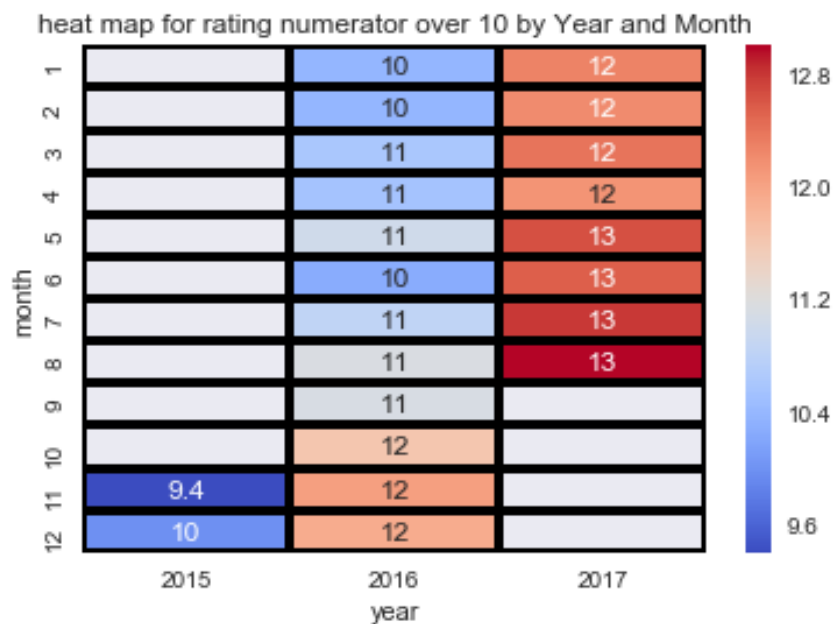
Follow

This is Logan, the Chow who lived. He solemnly swears he's up to lots of good. H\*ckin magical af 9.75/10



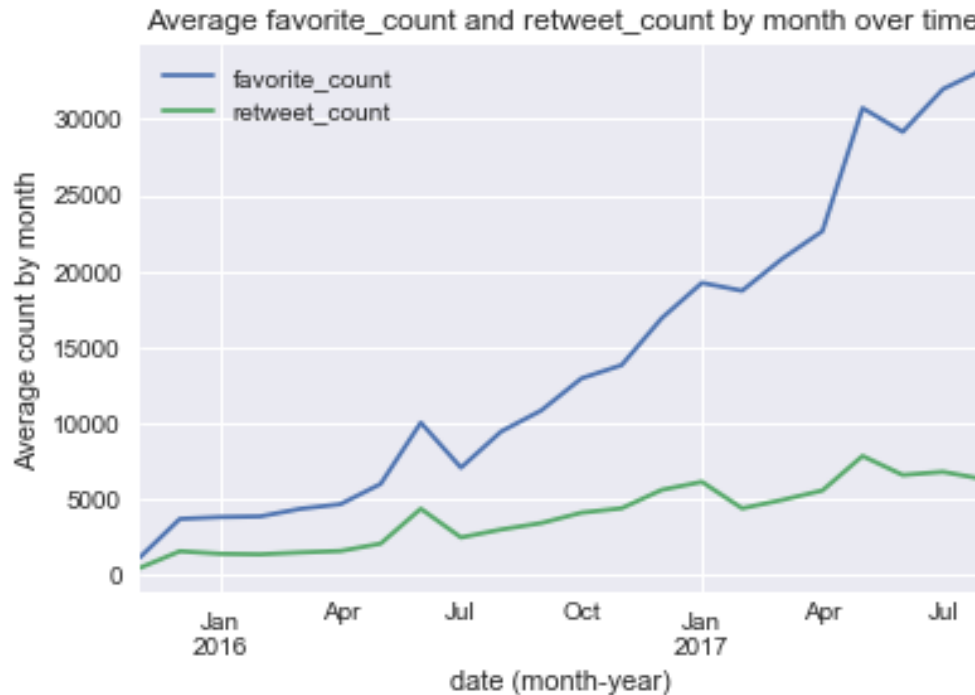
I am not sure if this spike in favorite count of around 20000! has something to do with the subject of the tweet here Logan/Harry potter. This what makes WeRateDogs so special is the amount of interaction between fans and the tweets depending on the subject or scene.

Now going back to the special rating methodology, I wanted to explore how the rating has evolved over time, using this heat map, below:

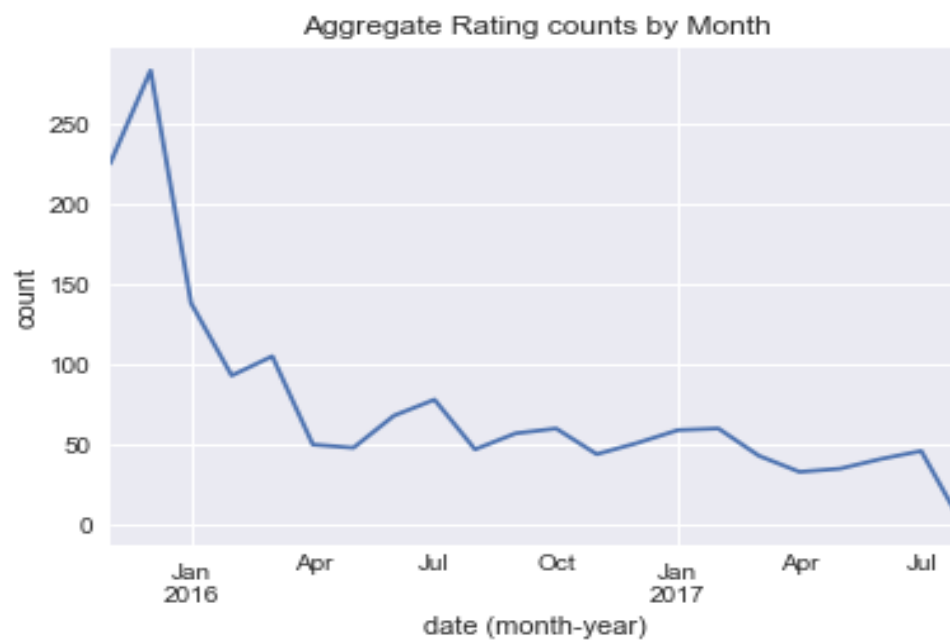


It appears that the rating numerator has increased from 2015 up the August 2017. Well I am not sure if the standard of the rating system have changed over time, interesting enough this has sparked a debate back in September 2016, link: <https://www.esquire.com/lifestyle/a54940/we-rate-dogs-matt-nelson-interview/>.

And sure enough, from the plot below we clearly see that the average favorite count showing an upward trend overtime



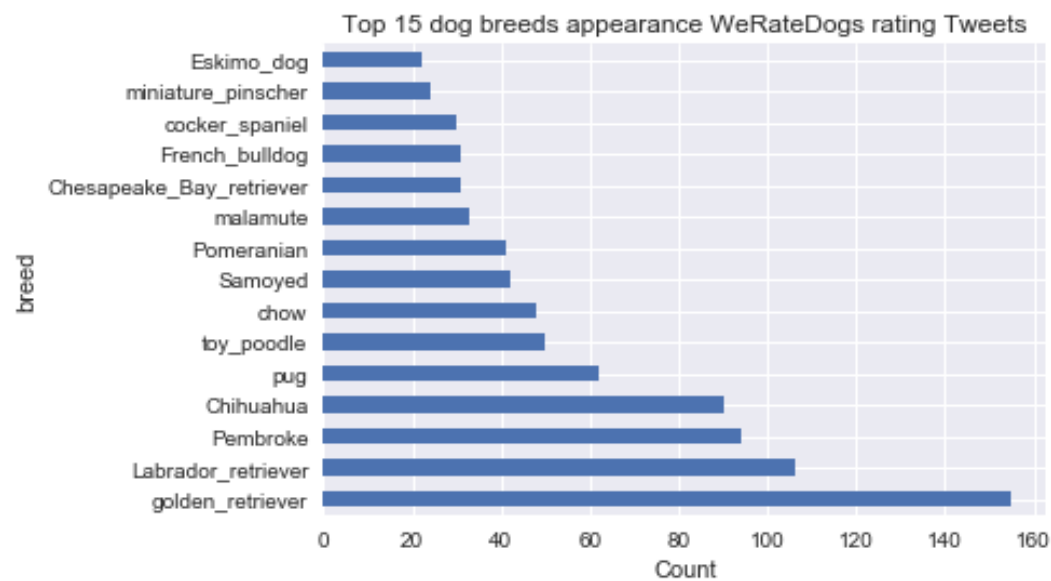
While the Rating tweets count per month appears to be in decline over time:



Moving to dog stages, though there were few observations where we were able to identify the dog stage from the tweet, but it is OK to mention that dog stage 'puppo' appears to have gained the highest Favorite count (computed in Median)

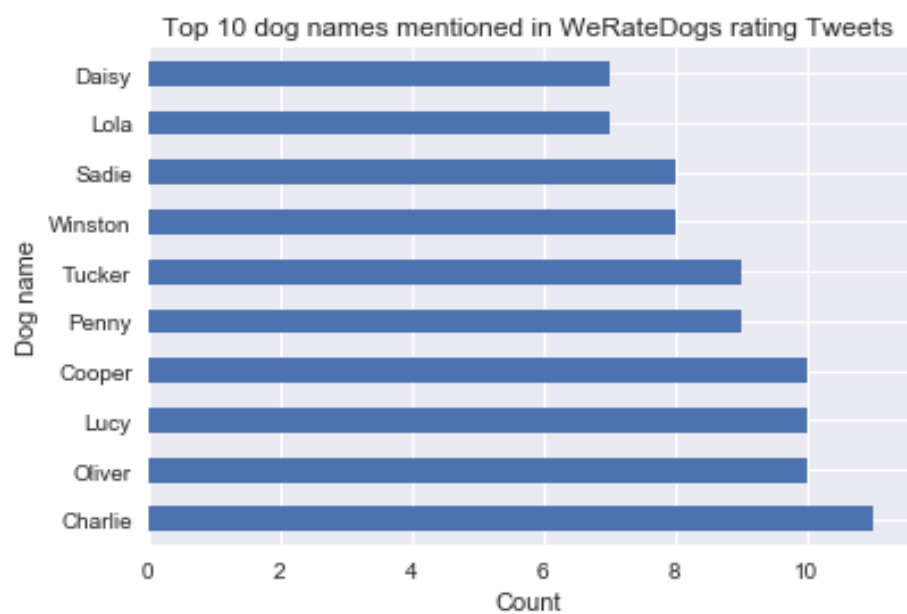
dog_stage	Favorite count (median)	count
puppo	13143.0	22
doggo	11701.5	54
floofer	11628.5	8
pupper	3283.0	168

Grouping by dog breeds, here are the top 15 dog breeds appearing in 'WeRateDogs' rating Tweets



With our winner the 'Golden Retriever' with more than 150 Tweets. Not quite surprising for me though as this is one of the few dog breeds that I actually know.

Finally we can check below the top 10 Dog names mentioned by WeRateDogs Tweets. This data was extracted from 1372 Tweets' text, where actually the dog name was mentioned.



## Conclusion

It appears that there is a moderate correlation between the rating and the favorite count, but rating alone is not enough to explain the change in favorite count as other factors should be taken into consideration like people opinions, whether the tweet includes a trending Hashtag or not, or if the image relate to important dates or maybe promoting good cause.