Blog post

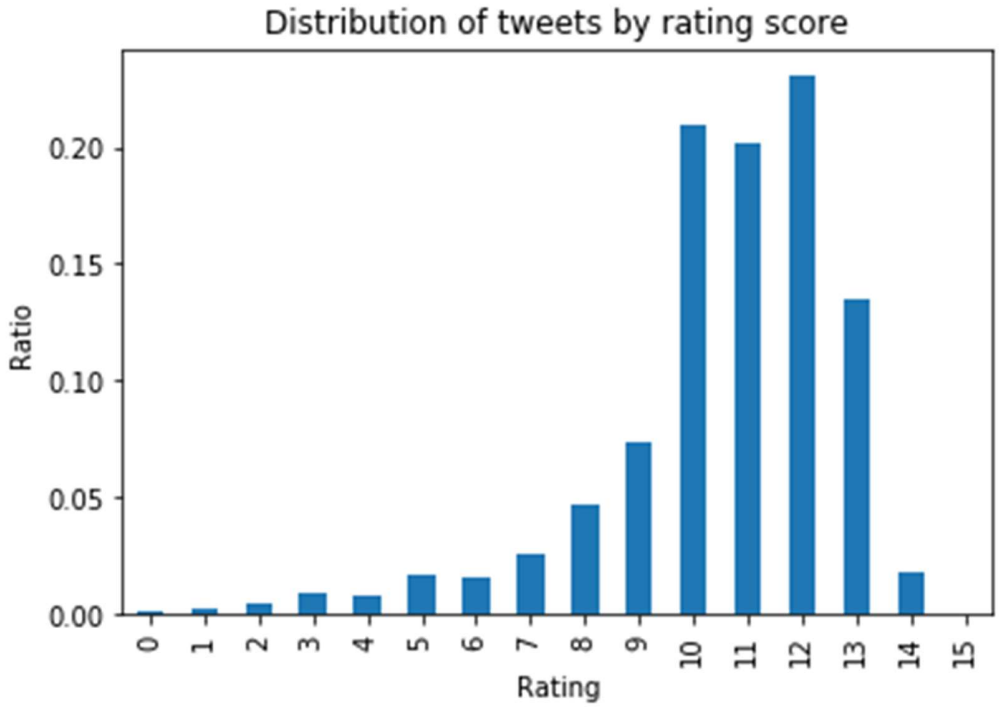
“They’re all good dogs Bret” triggers ratings inflation

If you’re a WeRateDogs fan you probably already know. What used to be a 9/10 or 10/10 rating on the cuteness scale for a dog or a puppy (or doggo or pupper), or even a very logical(?) 11/10 or 12/10 rating, is shifting to the extreme.

From the 2016 memes mania of “They’re all good dogs Bret” onwards there seems to be more and more occurrences of 13/10 and even 14/10 ratings, some users have pointed out.

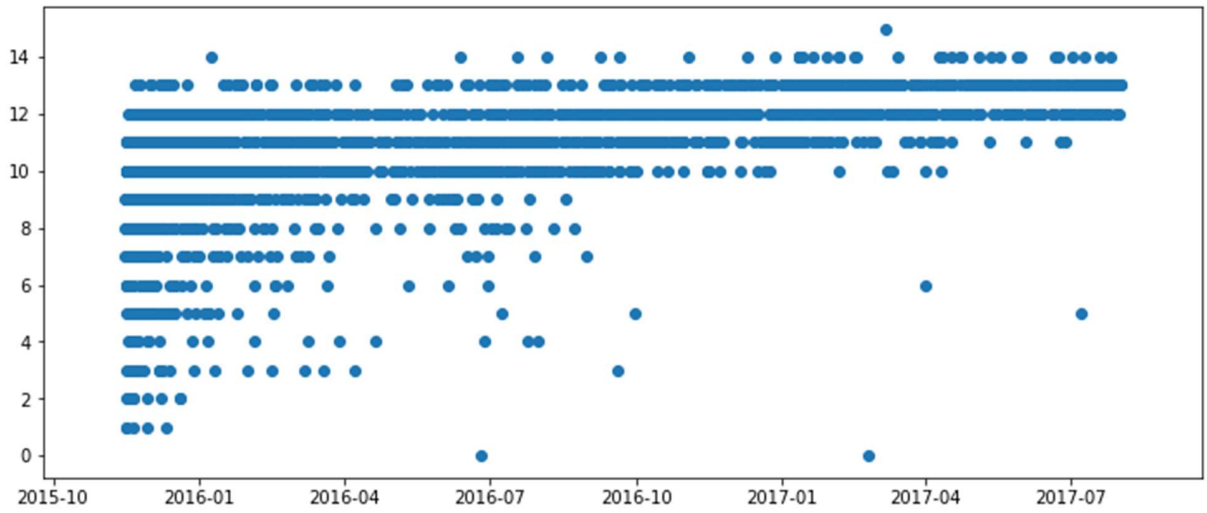
Now we have the data to confirm.

First let’s look at the relative ratings distribution for the last few years.



While most are 10 or higher, there are also some with lower ratings.

Now let's see how this has changed over time.



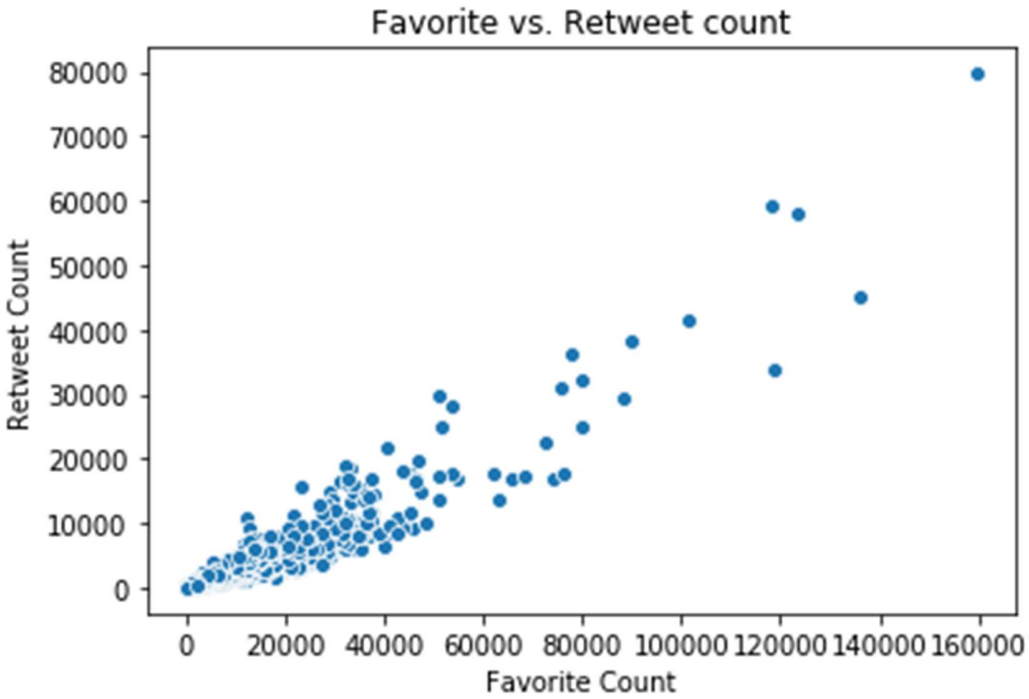
No doubt, there is a gradual, but visible shift during 2016, that the ratings below 10 disappear. As we move into 2017 there's also some clear "inflation" in the high end. For instance, the earlier common score of 10 gets less used, while the earlier very rare score of 14 becomes more normal.

The data has spoken:

"They're all good dogs Bret" did trigger ratings inflation! 😊

Let's look at some other plots, while we are at it.

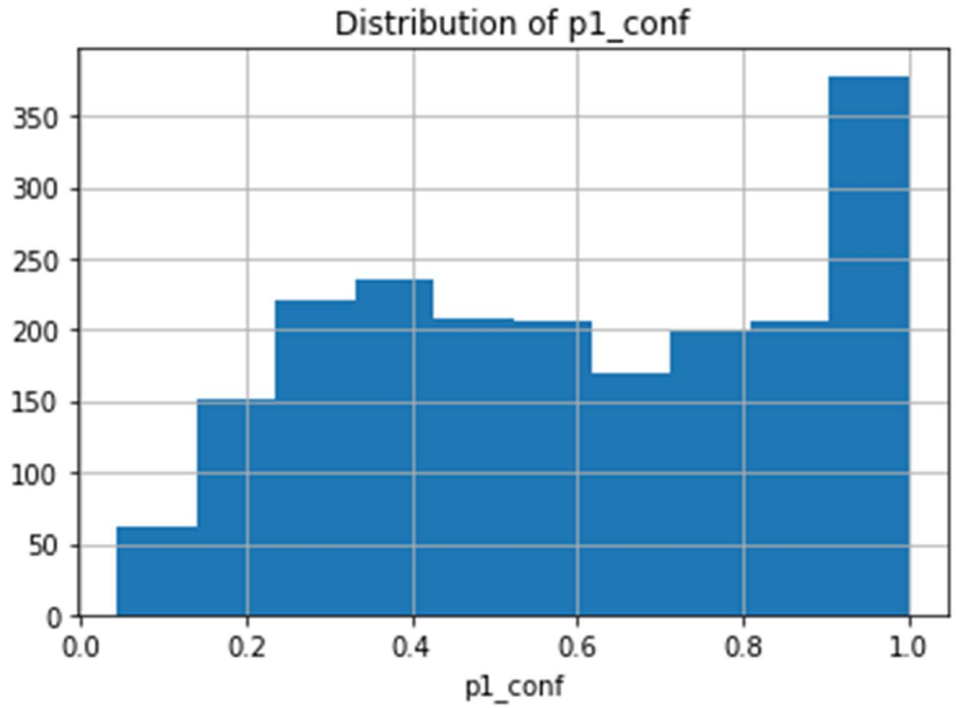
Is there any correlation between favorite and retweet count?



Yes. There is a moderate positive correlation.

Finally, we have got hold of some neural network modelled prediction data. One of the estimates “p1 confidence” says something about the likelihood that the image in question (the image with the highest likelihood if there are more than one image - could be up to 4) is a dog. Like a probability this number is between 0 and 1.

How the distribution of the p1 confidence score?



We can see from the histogram that the peak is near 1, ie. in many cases the model is very sure, say at least 90 percent sure, that the image in question is a dog. There is also a smaller peak around 35 percent. It is very rare that the model has a p1_conf of less than 10 percent.

In the 90+ percent sure case we can imagine the model only thinks the photo resembles a ‘dog’.

Or it could be a case where it might have two “main candidates” (def: non-main candidates would have less than 5 percent confidence), say ‘dog’ and ‘potato’.

Then again there could be three main candidates, say ‘watermelon’, ‘dog’ and ‘sofa’.

Etc.

Could this discrete number or “main candidates” explain the local peak in the distribution around 0.35-0.40? Please give your feedback in the comments section.