

Wrangle Act Report

After cleaning all the data we have, I have some interesting insights by analyzing the data.

First, there are 2097 original tweets which are focused on rating of dogs. These tweets have been post from as early as 2015-11-15 and latest in the data is on 2017-08-01. It is 625 days and the average number of tweets everyday is 3.3552. Actually I'm surprised by the number of tweets on each day. I thought the user would post more tweets since it is a little bit more like a company run twitter account now. And I believe most people like me would want to see more cute dogs.

Second, I looked at the favorite count of tweets. The maximum favorite count is 131820 and retweet count is 78754, and the minimum for both is 0. That's a large number of favorite counts and retweet counts. Also interestingly we can find even though lots of people like to favorite a tweet instead of retweet it. The minimum number for favorite and retweet count are both 0, which means some tweets are probably just retweets or replies, and no user will retweet it or favorite it.

Third, by looking at the mean and median number of favorite count and retweets count, we can find mean for both are 8110 and 3097, and median for both are 3585 and 1445. All those numbers are large, even though compared with tweets from some celebrities this is not significant, but it still shows WeRateDogs has a large amount of followers and people love its tweets.

Last, I have boxplots for both favorite count and retweets count, so we can have a more straightforward feel of the data.

These are some very interesting finds and similar techniques can be applied to analyze other users' twitter account, which is worth a try.

