

## **The WeRateDogs Project**

### ***An analysis Report of Data wrangled from Twitter***

We RateDogs is a twitter account that, as its name implies, it rates dogs! It receives dog pictures, it posts and comments them and provides a rating for the dogs. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. It has a very large following on Twitter with more than 4,8 million followers. It's comments are humorous and it has received international media coverage.

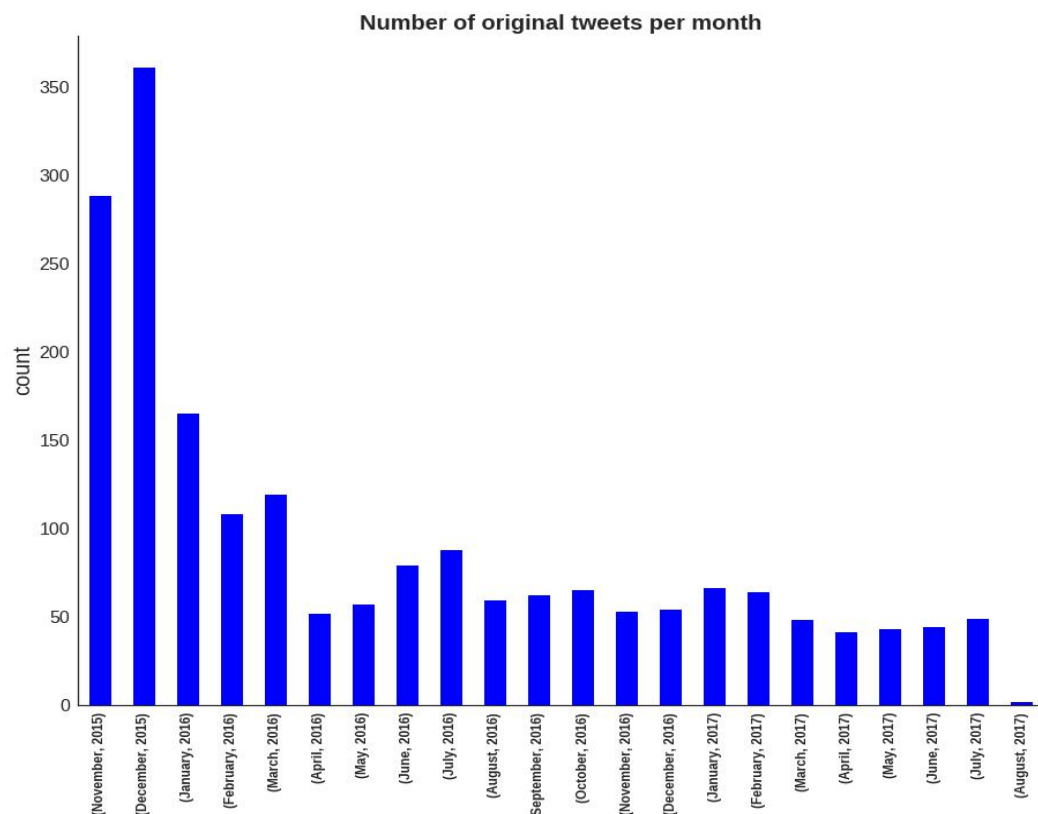


A typical WeRateDogs tweet...

As part of the Data Analyst Nanodegree offered by Udacity (that I am currently enrolled on) we were asked to analyse the twitter archive of WeRateDogs. WeRateDogs downloaded their Twitter archive and sent it to Udacity. This archive contained basic tweet data (tweet ID, timestamp, text, etc.) for all 5000+ of their tweets as they stood on August 1, 2017. I spent a great deal of time trying to clean the dataset, using the Twitter API to extract additional info for the tweets and finally analysing the dataset. This report presents some of the insights found in the dataset and displays the visualizations produced from the data.

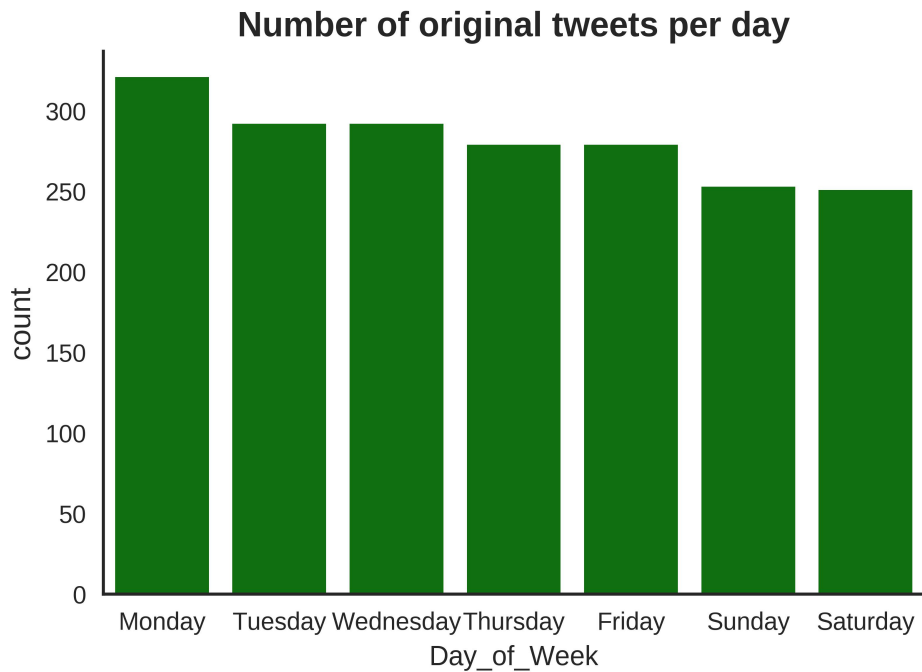
The final dataset had 1967 tweets from 15 November of 2015 until 1 of August of 2017. On average each dog got a rating of 12.3 and it was favorite by 8,959 twitter users. Respectively, the retweet count was 2,771. The three most common dog ratings (ignoring the denominator) were 12, 10 and 11. The most liked tweet gathered 131,813 likes and the most retweeted one 78,740 retweets. The most common dog name was Charlie with 11 appearances in the dataset, followed by Oliver, Cooper and Lucy (10 times each).

But lets check some statistics about the WeRateDogs twitter per se.

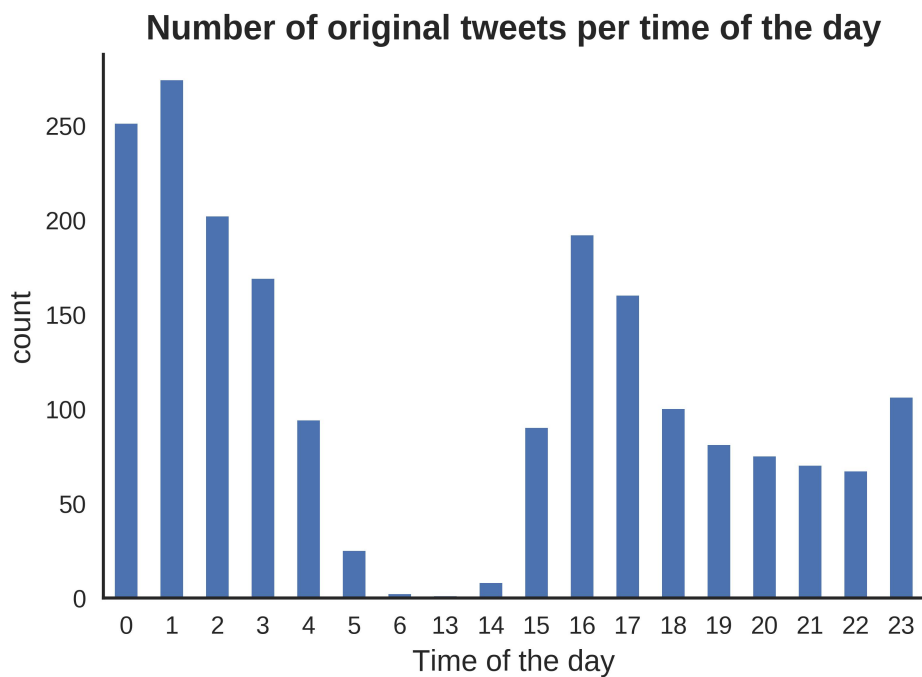


In the above graph we can see that November and December of 2015 were the most busy months for WeRateDogs twitter account with the number of original posted tweets exceeding 250 and 350 respectively. This is probably due to the fact that WeRateDogs was trying to increase rapidly its number of followers. From April 2016 and onwards no month had more than 100 tweets posted.

The graph below reveals that WeRateDogs was more active on Mondays and less active on Saturdays. There is a decreasing trend in the number of tweets posted as we approach the end of the week.



With so many followers probably the moderator of WeRateDogs account spent many hours during the day posting tweets.



In general the account holder of WeRateDogs seems to be active from 3 p.m. until 5 a.m. every day (14 hours per day!). He also prefers tweeting from an iPhone (98% of the time).

As a final note let's check which tweet was the most retweeted one: <https://t.co/7wE9LTEXC4> It shows a dog in a pool realizing that it can stand above the water even if he doesn't swim... It was retweeted 78,740 times.