

Act Report

Marc Campmany

March 2021

Introduction

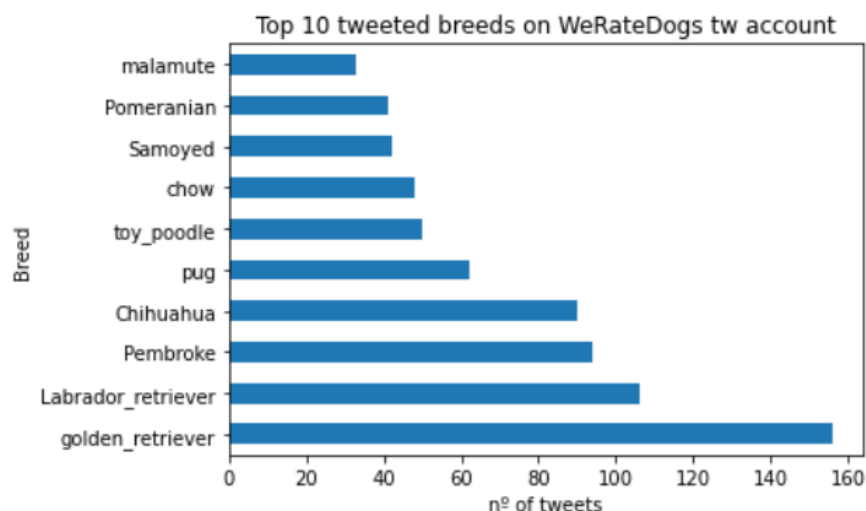
The dataset used for this project is the tweet archive of Twitter user @dog_rates, also known as WeRateDogs. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent." WeRateDogs has over 4 million followers and has received international media coverage.

Analysis and visualizations

The questions I want to answer are the following:

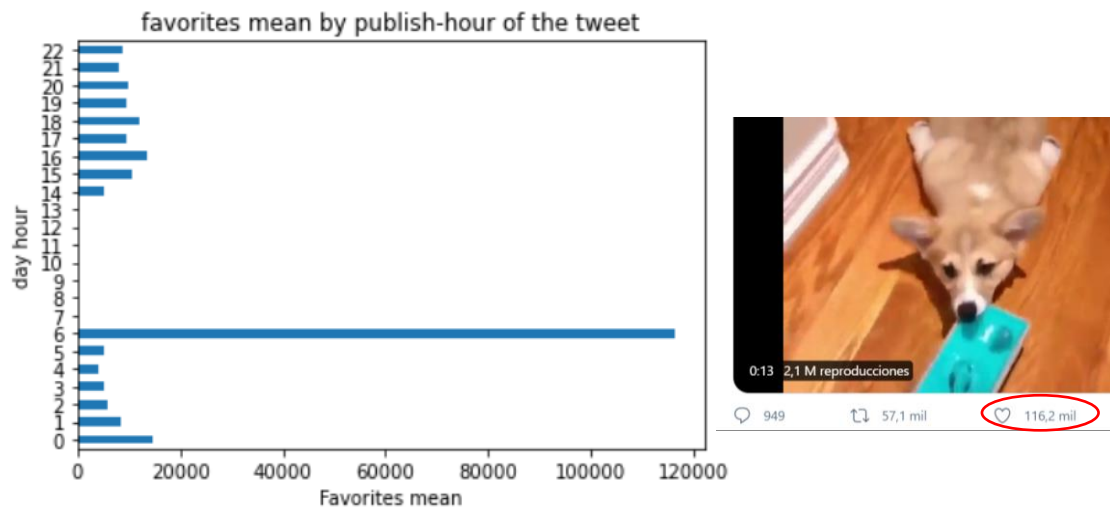
1. Which are the top tweeted breeds
2. Is the tweeting hour related with the retweets and likes obtained?
3. Breeds with highest ratings

1. Which are the top tweeted breeds

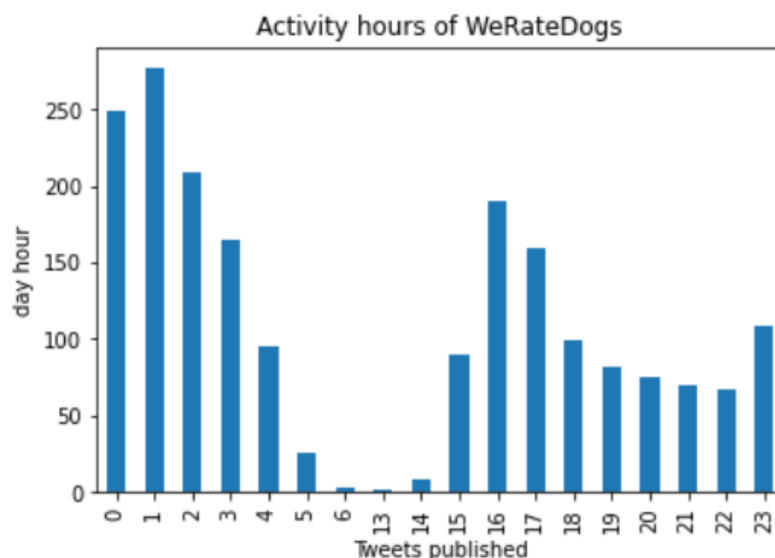


The most tweeted breed is golden retriever.

2. Is the tweeting hour related with the retweets and likes obtained?

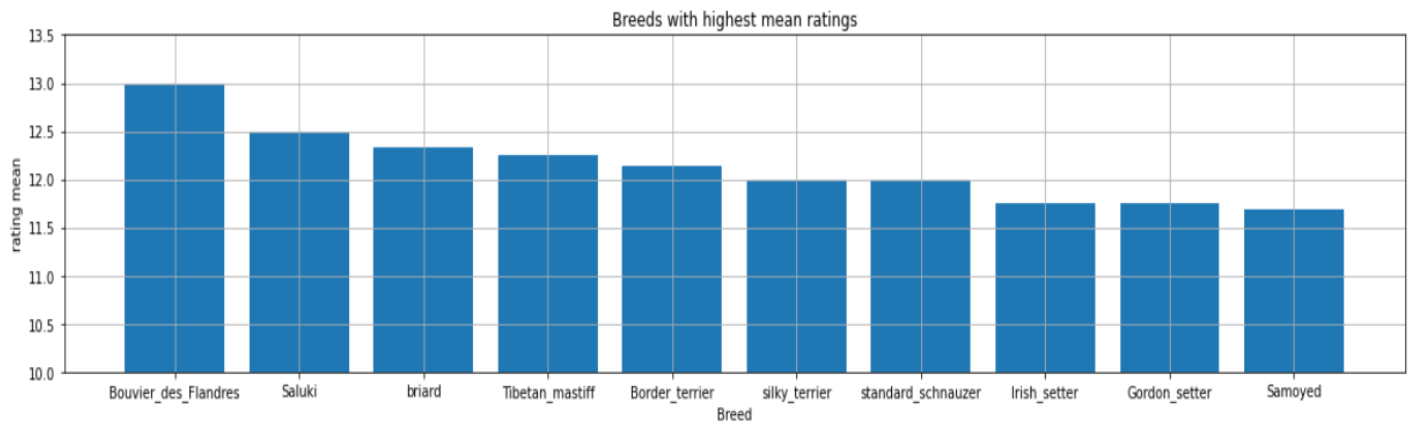


2.2 Activity hours of WeRateDogs

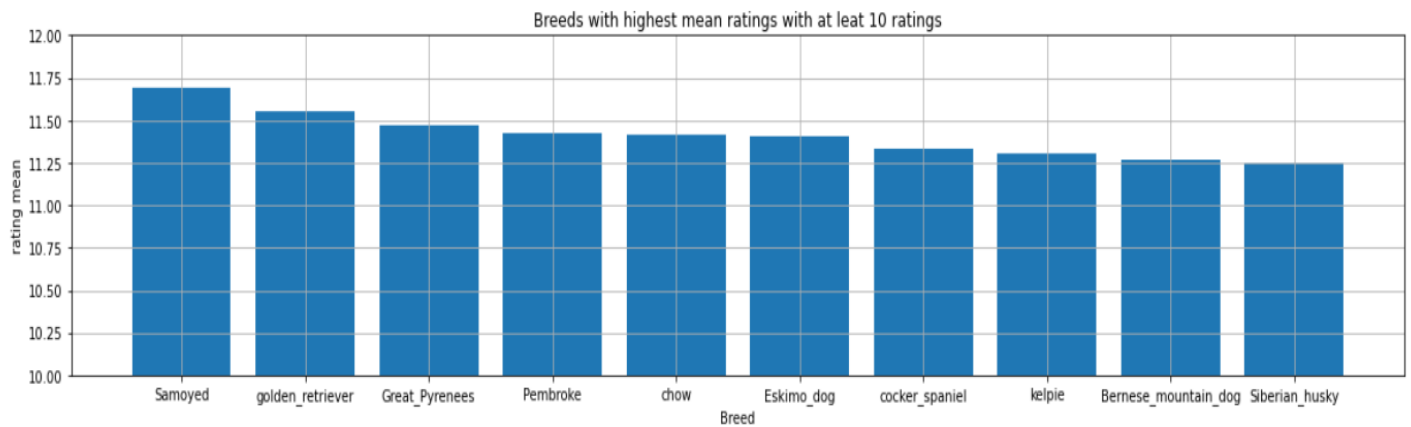


We can see that the hours with more activity are late at night or at evening. I suspect that these hours are translated to my time zone as it is strange to have so many tweets between 2-4 in the morning.

3. Breeds with highest ratings



The top rated breed by mean is "Bouvier des Flandres" with a rating =13 but with only 1 rating. We'll do the same analysis but with the condition of having at least 10 tweets with likes available.



We can now say that with this plot, the mean does not come from a single tweet and that the highest rated breed is "Samoyed" followed by "Golden retriever".