Findings Report - WeRateDogs Twitter Page

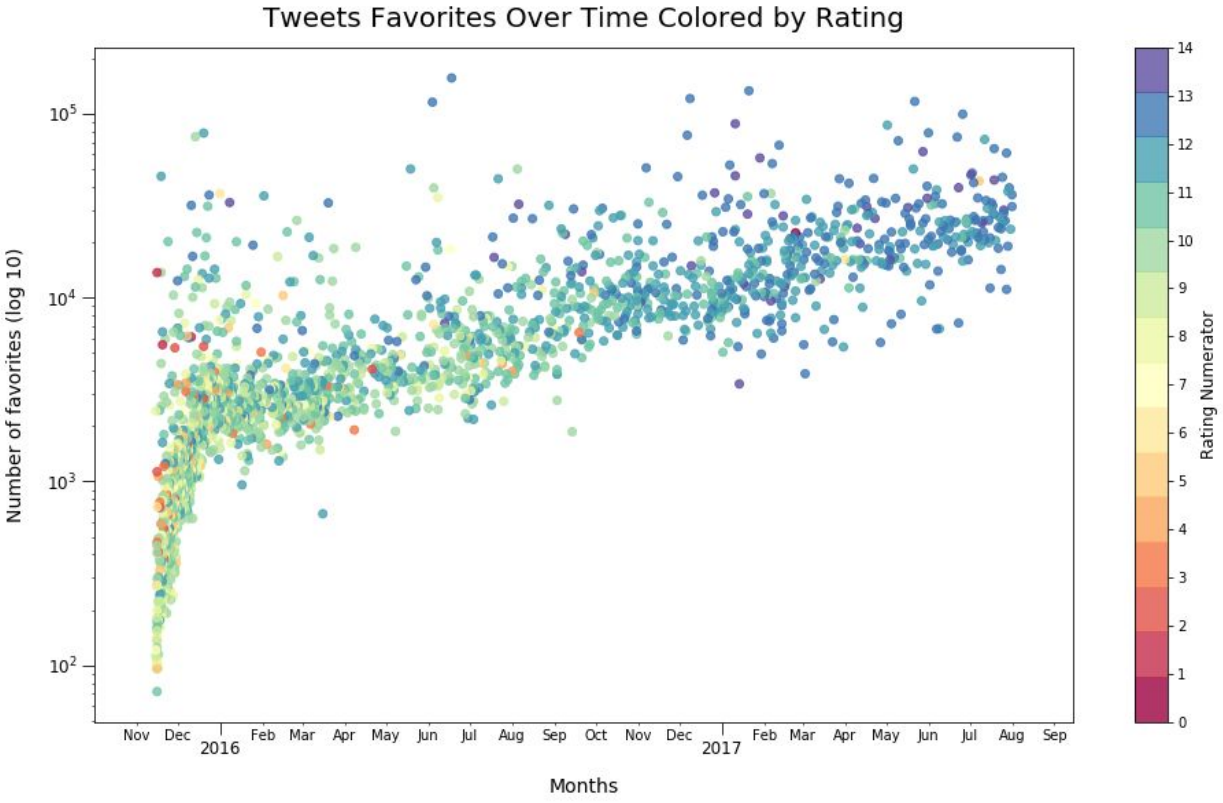
Gabriel Medeiros das Neves
May 2020

The main purpose of this report is to communicate insights found during the analysis process for the WeRateDogs Twitter Page, as a part of the Udacity's Data Analyst Nanodegree.

1. Ratings inconsistencies over time

Although the page has maintained the same rating system over time (with the exception of some intentionally non-standard tweets), it was possible to realize that the ratings gradually increased after 2015.

Besides that, dogs with low ratings, which were already rare at the beginning of the page, became almost nonexistent in 2017, as visually shown below:



2. Changes in publication hours over time

Despite it was not possible to find significant correlations between the time of publication and the number of favorites, it was noticed that the publication times changed after about July 2016, when the page started to post more frequently in the afternoon and late afternoon, unlike the previous tweets, that were mostly published in the early hours of the day (between midnight and 6 AM).

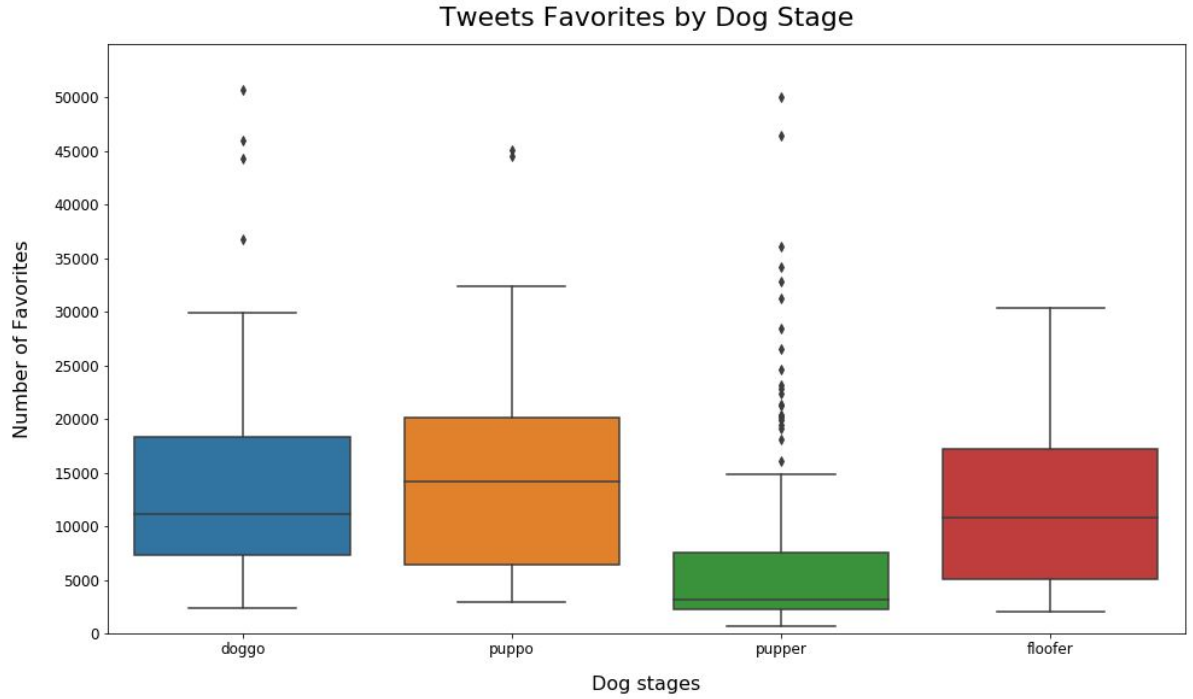
There are several reasons why this change may happened, but it does not appear to have any negative effects and is recommended since many users are already asleep after 2 AM and may not see the tweet the next day in their timelines.

2. Preferences for publications of certain dog stages

Before anything, it's important to note that the dogs' stages were informed in just 303 tweets, which complicates the formulation of conclusions and recommendations, since the tweets may have been posted at different times and contexts. That being said, I'll be describing the findings below.

When analyzing the relationship between the number of favorites and the dog stages, I was surprised that the puppers were considerably behind the other stages, since it's hard to find someone that doesn't like a pupper. However, there were many tweets about *puppers* that reached a high number of favorites, which leads me to believe that they are still very much loved and that their average may have been a little lower because many of their tweets were published in the beginning of the page or didn't have the best image or caption.

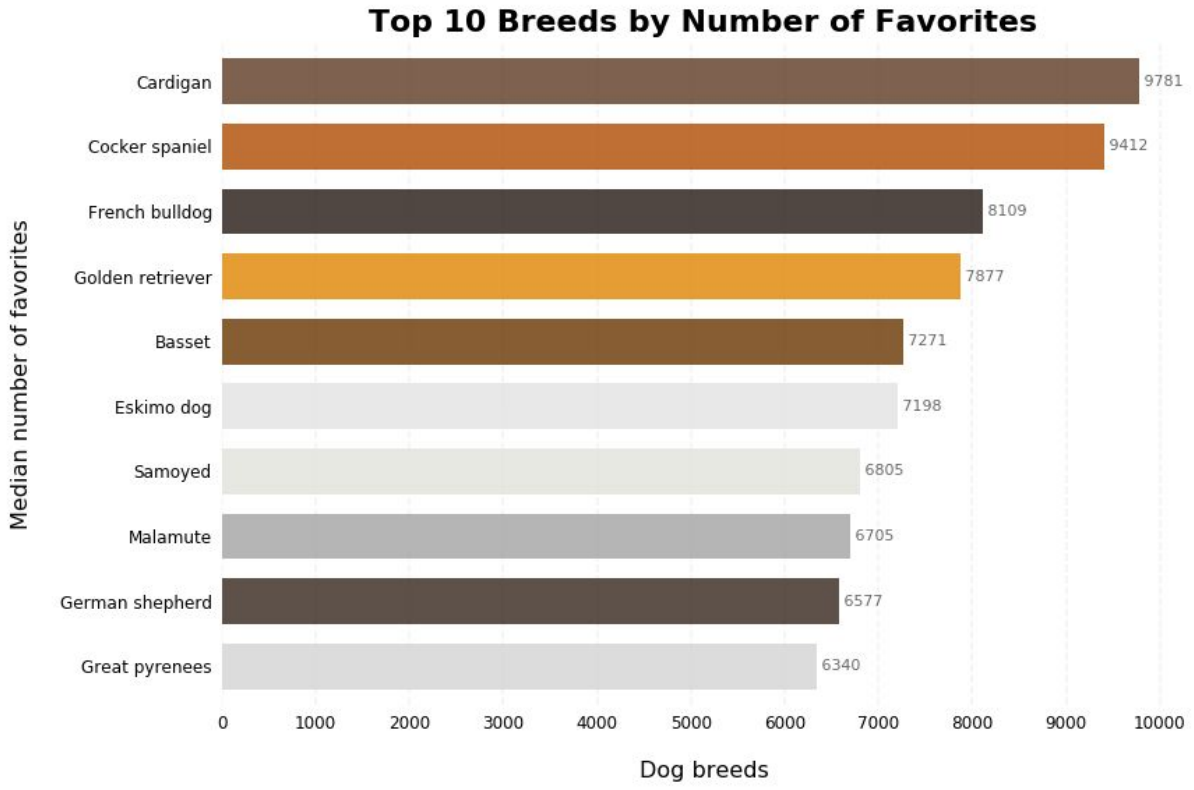
On the other hand, *puppos*, *doggos* and *flooflers* proved to be good dogs and received lots of "little hearts" from the public in their tweets (with *puppos* being the most popular stage), as shown in the following box plot:



4. Popular breeds

Each breed of dog has its physical and behavioral characteristics that differ from all others. These characteristics do not make a dog good or bad, at most they make a dog look like it would wear a suit or some type of colorful shorts. Anyhow, the public apparently has their favorites, such as the *cardigan* and *cocker spaniel* breeds, which have shown the highest level of public acceptance.

To measure the acceptance of the dog breeds, I calculated the median number of favorites for each breed that was identified in 10 tweets or more and created the following bar chart:



5. Recommendations

The high ratings do not seem to hinder the growth of the page in any way and have become a brand after the "They are good dogs Brent" meme, so there is no need to change the rating system.

Regarding the time of the posts, though there is no significant correlation, it's probably a good idea to keep the publish time out of the early hours of the day (after all, "nothing good happens after 2 AM"), since many people are already sleeping at that time and may miss the tweet the next day.

Furthermore, concerning the dog breeds, as textually and visually described in the section above, some breeds have a better response from the public. This does not mean that only publications with the breeds present in the Top 10 Chart should be made, seeing that the diversity of the dogs is a very important factor for the page popularity, but that some breeds can appear more often than others, since they are more popular.

Similarly with dog breeds, page followers also seem to have a preference for a specific stage: puppos. Perhaps it would be interesting to conduct tests to confirm that puppos tweets really receive more favorites, since it's difficult to confirm correlations when there's multiple variables associated with the success metric. In addition, it would be really useful for future analyzes if the dogs' stages were informed more frequently.