

Udacity Data Analyst Nanodegree - Project 4: Wrangle and analyze data - Vasileios Garyfallos - April 2020

WeRateDogs

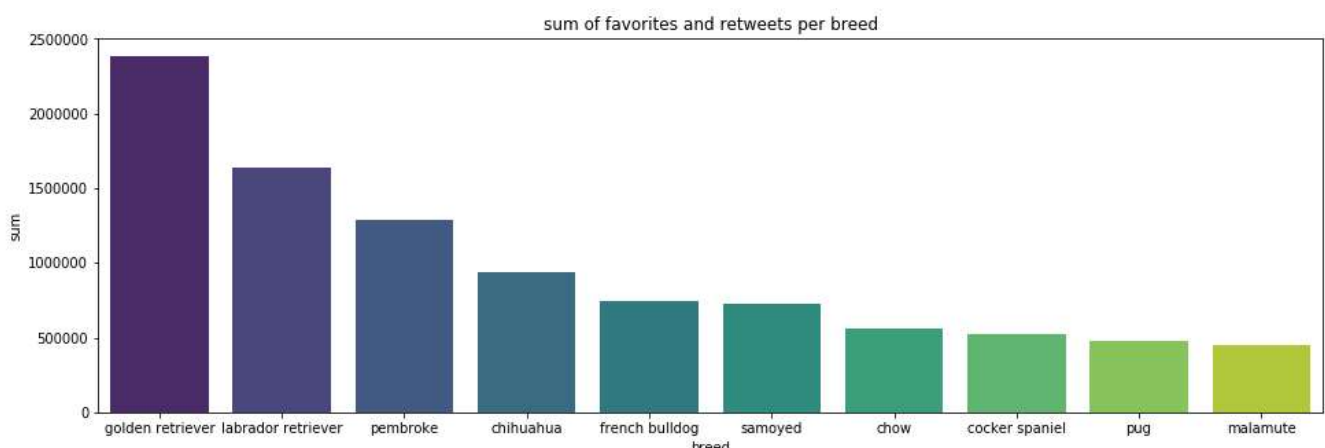
With over 8.7 million followers as of April 2020, including J.K. Rowling and Lin-Manuel Miranda, WeRateDogs' absurdist content clearly enjoys mass appeal. The account, started by Matt Nelson in November 2015, was inspired by so-called "Weird Twitter", a broad and loosely associated group of Twitter users known for their eclectic and amorphous humor.

Users submit photos of their dogs, which, if selected, are later tweeted along with captions and ratings. While the ratings are ostensibly on a scale of one-to-ten, most dogs receive ratings like "12/10". Although these ratings are one of the account's most characteristic features (spawning the viral "They're good dogs, Brent" meme), the extremity of its ratings has decreased since the blog first began.

Questions, Answers & Visualisations

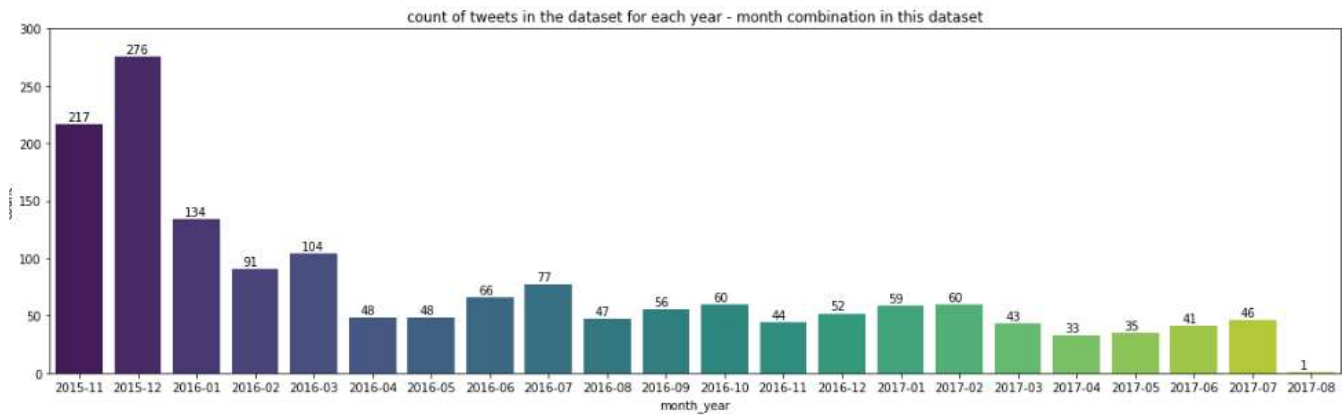
Question 1: Based on the predicted, most likely dog breed: Which breed gets retweeted and favorited the most overall?

As we can see in the graph below, Golden Retriever is the breed who gets most retweeted and is the most favorited among all breeds with Labrador Retriever following with significantly less favorites and retweets over the examined timeframe.

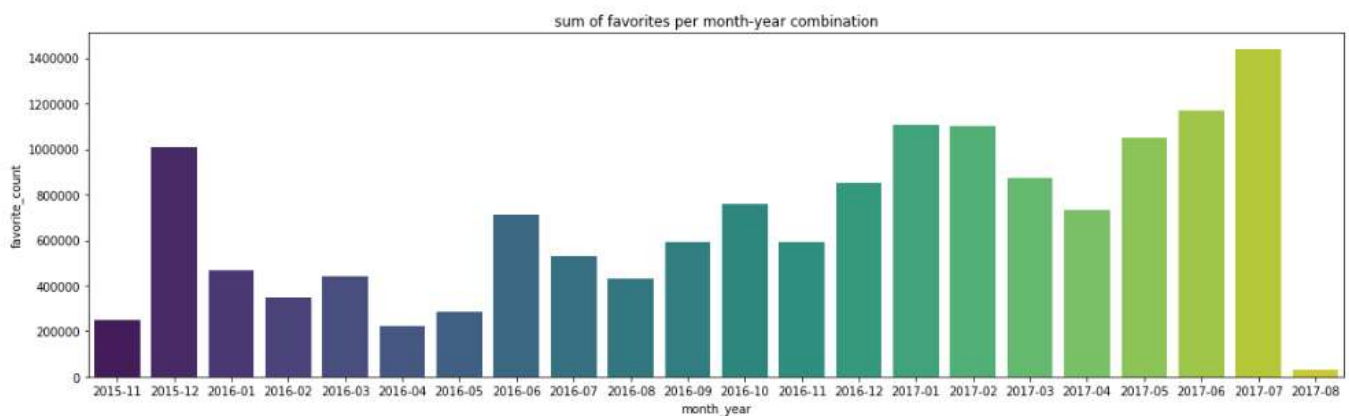


Question 2: How did the account develop (number of tweets, retweets, favorites, image number and length of the tweets)?

Over the whole timeperiod the number of monthly tweets of The WeRateDogs account has seen a decrease. In April 2016 the number of tweets rapidly dropped and since then it has a relatively stable level to lower levels.



Interestingly enough, there is higher engagement of users when seeing WeRateDogs content in their feed, regardless the fact that the overall numbers of tweets has seen a decrease. We have less content shared from the account, but the content engages more users reactions. The graph below clearly depicts this:



Conclusively, we can say that there is a negative correlation between number of tweets - engagement of the tweets over the examined timeframe. This indicates a very good content strategy from the account owners over the timeline. Quality of the content always surpasses quantity of the content.

Question 3: Is there a pattern visible in the timing of the tweets?

In years 2015 & 2016, the most posts are during the night between 0:00 - 5:00. Between 4:00 - 15:00 there is a very small amount of tweets, no tweets between 7:00 - 12:00 and at 14:00. There are a few tweets after 14:00, but not as many as between 0:00 - 5:00.

See the graphs below:

