

# act\_report

August 10, 2020

## 1 Act Report

### 1.1 Overview

The `twitter_archive_master` dataset contains tweet data for 1964 tweets from the @dog\_rates handle on Twitter 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.

In this report, I present some insights and visualization from the analysis of this dataset.

### 1.2 The Dataset

The `twitter_archive_master` dataset contains data that has been extracted from the *WeRateDogs* account such as the Tweet's ID, when it was posted, the source via which it was posted, and the Tweet's text along with the number of favorites and retweets. It also contains information regarding the dog tweeted about: name, the url for the dog's picture, the rating, and the stage of the dog (i.e. doggo, pupper, puppo or floofer). In addition to all the previous information, the results of the prediction of a neural network that classifies the breed of the dog is also included in the data.

The `twitter_archive_master` dataset is the result of an intensive labor of data wrangling. The data was gathered from three different sources, assessed for quality and tidiness issues, and then underwent an extensive cleaning procedure.

### 1.3 The Analysis

Through the dataset, I managed to answer three questions:

**1. Which dog breed has the highest retweets/favorites?** To answer this question, I aggregated the number of retweets/favorites for every breed, sorted them and extracted the 10 breeds with the greatest number of retweets/favorites. I then did a bar plot for the retweets and favorites separately using the package *Plotly*.

The most retweeted breed is the Labrador Retriever followed by the Golden Retriever which is no surprise since the Retriever is the most popular breed of dog in many countries around the world .

In line with the results from the retweets figure, the Labrador Retriever and the Golden Retriever are the most favorited breed of dogs. We notice that the numbers of favorites exceeds the numbers of retweets which is common feature of Tweets on Twitter.

**2. Which dog stage has the highest retweets/favorites?** Similar to the procedure followed in the first question, I aggregated the number of retweets/favorites for every dog stage, sorted them and did a bar plot.

Doggos (older dogs) were the most retweeted.

However, puppies (younger dogs) were the most favorited.

**3. What does the rating of dogs look like?**

It is obvious that people love dogs too much that they actually rate them with scores greater than 10! Why? Again, because...

## **1.4 Insights**

From the previous analysis, one can extract the following insights: - People love Retrievers, retweeting and favoriting them the most of any other breed. - People retweet Doggos the most, but favorite Puppies more. - Favorite counts are always higher than retweet counts - The distribution of dog rating is skewed to the left, meaning that the majority of ratings are very high.