

WeRateDogs Twitter Archive: Data Analysis and Visualization

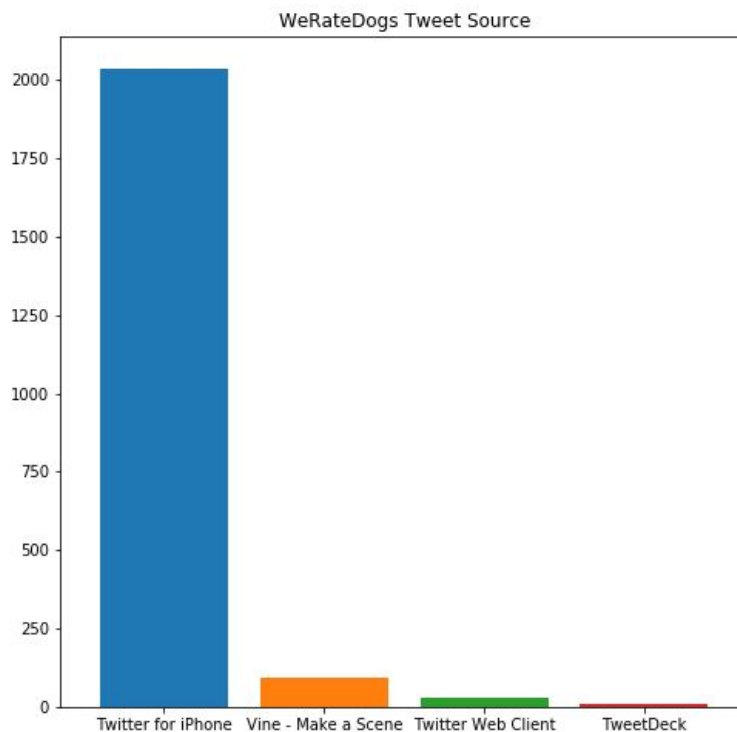
The Twitter account, WeRateDogs, has amassed a huge following of about 8.4M users (as of Sept 2019) in the past few years. It's no surprise, either. With pictures of dogs and puppies being shared everyday alongside quirky captions and off-the-chart "goodboy" ratings, this account can be appreciated by dog-lovers all around the world.

In this report, I will be sharing a few fun insights I gleaned by gathering, cleaning, and analyzing the tweet data from this account's Twitter archive.

*All data represents the WeRateDogs twitter archive between Dec 2015 and Aug 2017.

1. Tweet Source

All tweets in the WeRateDogs archive from the defined timeframe came from four different sources: iPhone, Vine, Twitter Web Client, and Tweetdeck. Of these different sources, the overwhelming majority of tweets came from iPhone. In fact, over 94% of 2,168 tweets were created from this source. Below, is a bar chart representing these proportions.



2. Image Prediction

A Udacity programmer ran the images from WeRateDogs' tweets through a neural network that can classify breeds of dogs. For each picture, there were three different predictions, p1, p2, and p3, in order of decreasing confidence levels. Sometimes, the program misidentified these pictures for non-dog things, such as random objects or other animals. I calculated the top 10 non-dog objects that were predicted at the p1 level. The results are shown in the table below.

Prediction at p1 level	# of times predicted
seat_belt	21
teddy	18
web_site	15
tennis_ball	9
dingo	9
hamster	7
doormat	7
Siamese_cat	7
bath_towel	7
swing	7

3. Retweet and Favorite Counts

I queried the most recent counts of "retweets" and "favorites" of WeRateDog's tweets and plotted a line graph of these counts against the timestamps of the tweets. The resulting plot shows the trend of the account's retweets and favorites over the 2015 to 2017 timeframe, which appears to be increasing over time, with favorites consistently outnumbering retweets. This trend aligns with the intuition that, over this 2-year span, the account gained more popularity and therefore was seen by more Twitter users and received more engagement. This increased engagement more notably pertains to favorites.

