

WeRateDogs Twitter Archive - Act Report

Project By Chizaram Emenyonu

The WeRateDogs Enhanced Twitter archive contains data extracted from 2356 of the 5000+ tweets from the @dog_rates twitter account, posted between November 15, 2015 and August 1, 2017. WeRateDogs developed its own dog classification, based on the dog's age and appearance. Puppies are "puppers", older puppies are "puppos", older puppies are "doggos" and hairy dogs (or almost every dog) are "floofers". The retweet count and favourite count for each tweet were not included in the enhanced archive, and so I had to programmatically download this additional data from the twitter account using the twitter API (Tweepy) from the archive. Along with the Twitter data, I also downloaded an image predictions file from Udacity servers containing the top 3 predictions for dog breeds based on the images from the tweets.

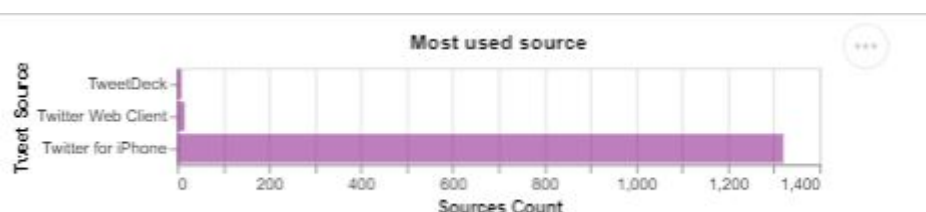
Wrangling Data

Before starting the analysis, I had to assess and clean the data to make it easier. I assessed the data both visually and programmatically for quality and tidiness and cleaned all errors I found. This assessment process was done over and over to achieve a level of cleanliness.

Insight

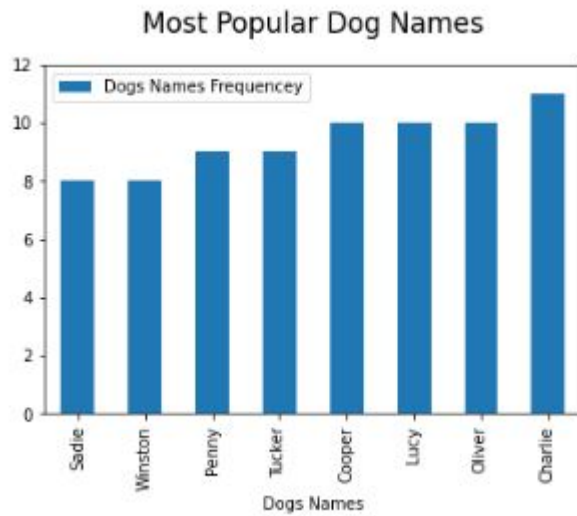
Twitter for iPhone is the most frequently used platform for posting tweets

One of the easiest insights I saw was that Twitter for iPhone app is the most used platform by the WeRateDogs twitter handle. About 98% of the tweets posted by WeRateDogs is from Twitter iPhone app. This is not really a surprise as the iPhone brand is one of the most popular phone brands in the world



Charlie is the most popular dog name

Grouping all the tweets by the dog name and running a count to find the most popular name, it showed that the most popular dog name. Cooper, Lucy and Oliver shared the second spot.



And the most loved dog is..... STEPHAN!!!

I decided to find out which dog had the highest numbers of likes and retweet. I was able to get this particular dog by running through the dataframe and filtering the favorite_count and retweet count columns to get the particular tweet with the highest count. Stephan received about 111230 favorite counts and 51443 retweets making him the most loved dog.