ANALYSIS OF DATA FROM WERATEDOGS TWITTER ARCHIVE

Twitter is a rich source of data. It is an application which has millions of active users interacting with one another every second of the day. That is why when we decided to practice some wrangling and analyzing of data, Twitter was a good choice.
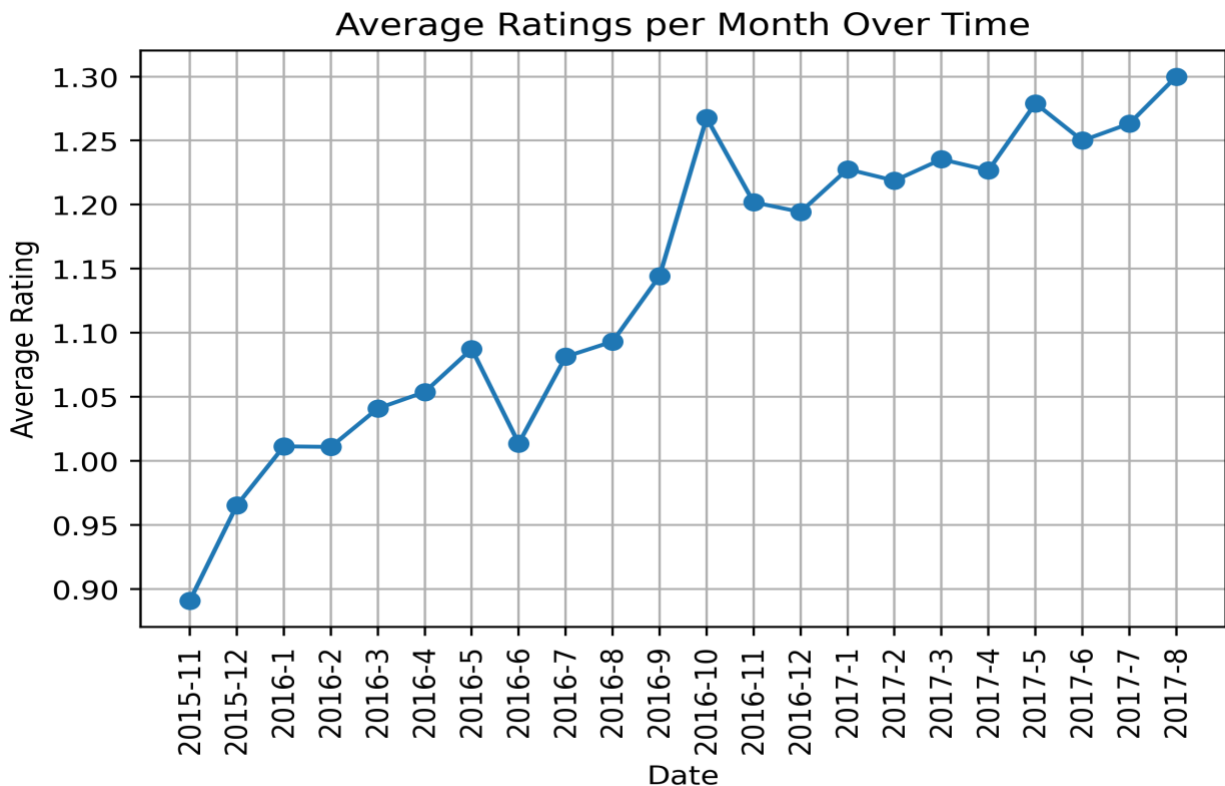
The main aim of this project was to analyze Tweet data from a handle called WeRateDogs. As the name implies, this account rates photos of cute dogs submitted to them by their owners. There were 3 datasets used: a provided archived dataset of tweets; a bunch of tweets and other features including retweets and favorite counts extracted from the Twitter API and a dataset containing model predictions of the dog breed submitted for rating.

After cleaning and analyzing we found the following interesting insights;
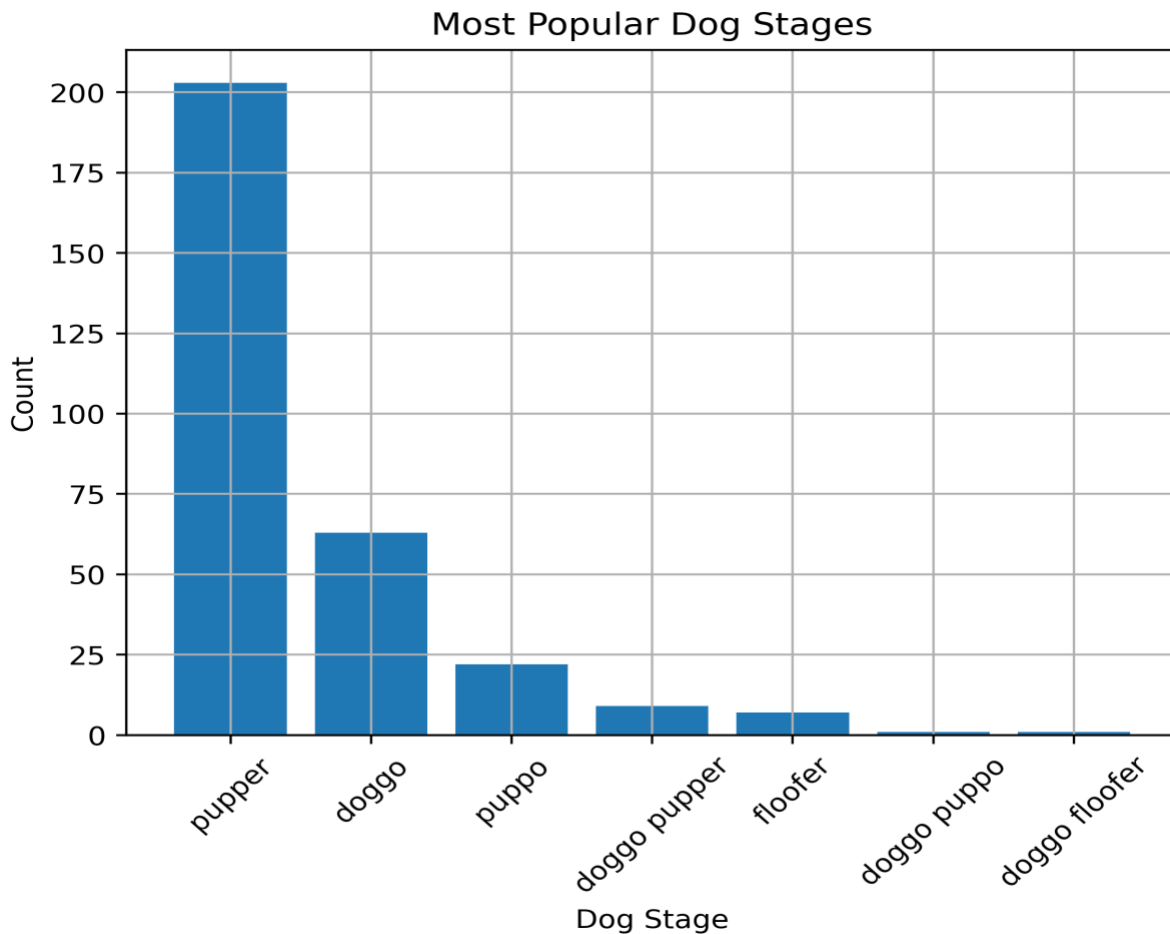
Ratings have gotten better over time.
We looked at the average rating score per month awarded to the dogs over the duration covered in the dataset i.e. November 2015 to August 2017 and discovered that the general trend was upwards (see figure below).
This could mean that the dogs have gotten cuter and cuter over time! It could also mean the adjudicators have become nicer but we prefer the former conclusion.
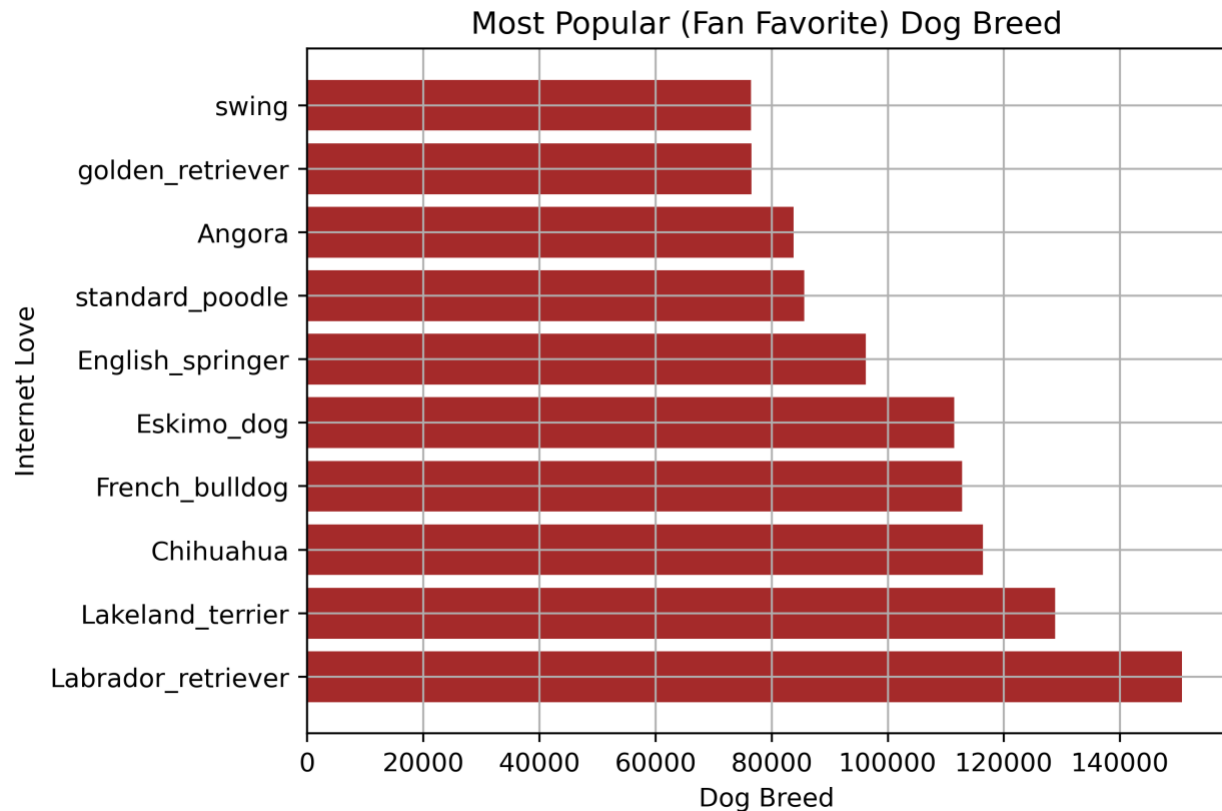
Puppers are the most popular dog stage.
There were 6 distinct dog stages in the data although the dog stages were missing for most of the data. Amongst those six, the most popular by a huge margin were Puppers (see figure below).

Labrador Retrievers are the favorites of the fans.
We ran the images of the dogs through a machine learning model to predict the names of the breed of the dogs as this information was not available from Twitter. We found that the Labrador Retriever was the darling of the fans with pictures of dogs of that breed amassing the highest number of Likes (over 150,000)!

Most Popular (Fan Favorite) Dog Breed

What popular name should give your dog so it can be loved by the fans?
We decided to check if there any names which appeared the most in the dataset, suggesting that they are popular dog names, which also seemed to have many Likes, suggesting that the fans love those dogs with those names as well.

We found that the intersection of these two sets included the following names: *Penny, Stanley, Bo and Charlie*. So, next time you want to name your dog, just pick one name from here and it is on its way to celebrity life.