

# WE RATE DOGS

Data Wrangling Project - Udacity

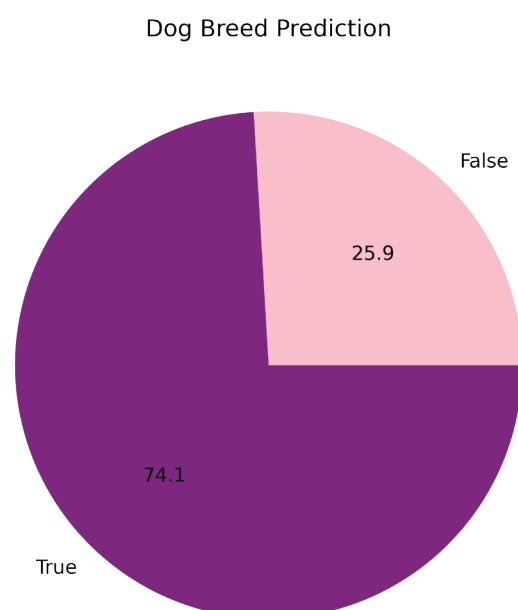
- Insights Derived From Analysis and Visualization

WeRateDog is a Twitter account comprising of people's ratings about their dogs, humorous comments, and images of the dogs. The data were gathered from three sources, cleaned, and visualized.

During analysis, the following insights were drawn:

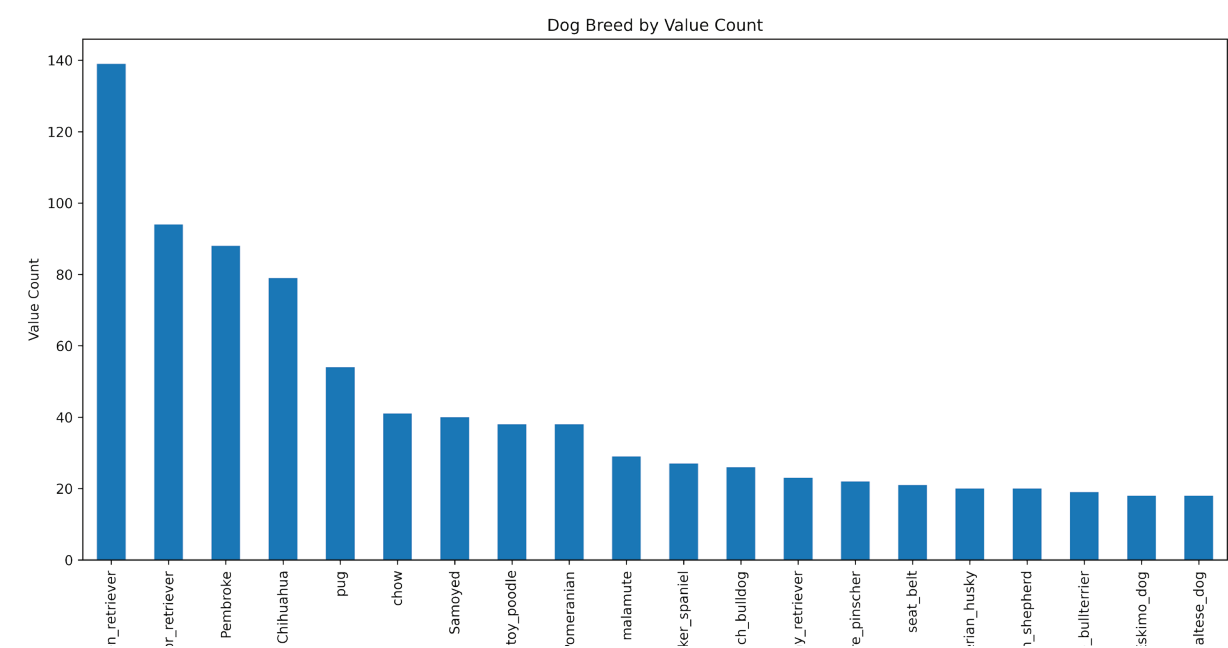
1. There are four dog stage category namely pupper, doggo, puppo and floofer. Some images were seen to have more than one dog belonging to the puppo and doggo dog category. Most dogs were categorised as pupper.
2. Amongst four source of tweets, Twitter for iphone is the most common source used by users to make a post.
3. A Labrodor retriever without a name and a rating of 13/10 had the highest number of retweet count. This tweet was posted on 18th June 2016.
4. Year 2016 recorded the highest number of retweets of WeRateDog tweets.
5. After cleaning the data, 1,161 dogs had ratings greater than 10.

From 1,993 posts obtained after cleaning, 74.1% were predicted to be dogs while 25.9% are not dog images. Some people posted images that are not dogs. While it was observed that some images had dogs in them but were classified wrongly due to other images in the background.

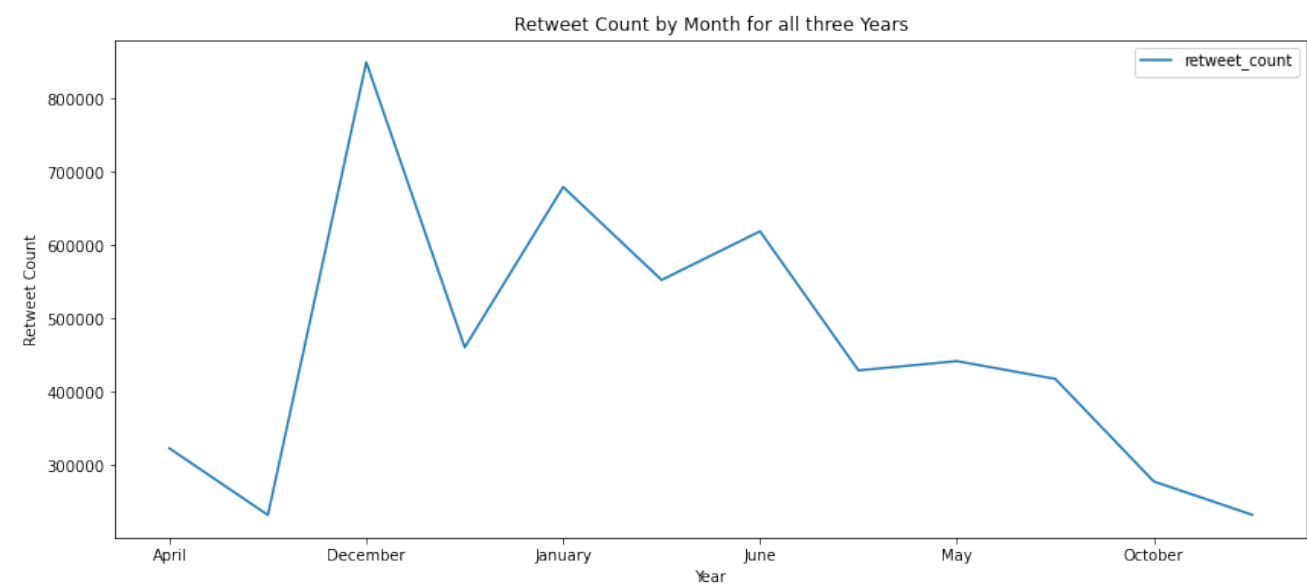


# Insight drawn from Visualization

Golden Retriever dog breed had the highest count in the dataset followed by Labrador\_retriever and pembroke. After prediction, about 139 images belonged to the golden\_retriever dog breed.



December, January and June had retweet counts greater than 600000 over the years



## Insight drawn from Visualization

A strong positive correlation exists between favorite count and retweet count.

