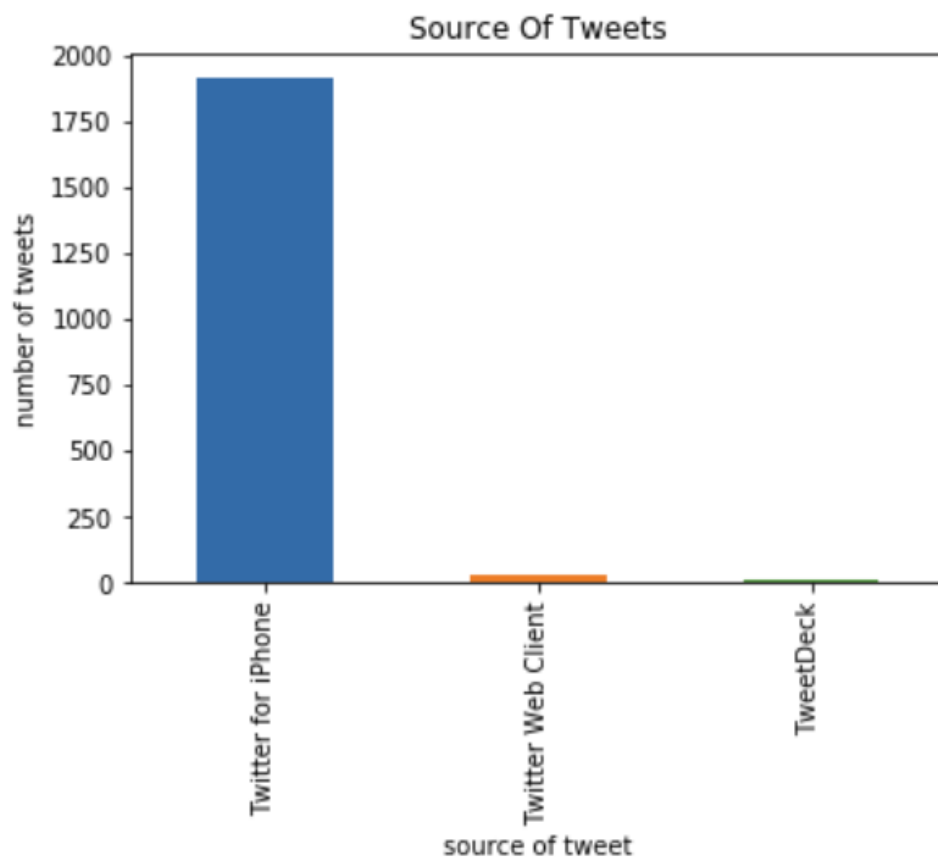# Act Report

By: Hassan Almaghrabi

Introduction

After I have finished the wrangling process and combined the clean data into one master dataframe, I started to set questions and analyze the data to figure out the answer to each question to come up with meaningful insights from the data, the questions that I stated are as follows:

Question 1: What are the most 5 dog breeds that are predicted by the model and how many distinct breeds are in the dataset?

There are 113 different dog breeds in the dataset and the most 5 breeds are:
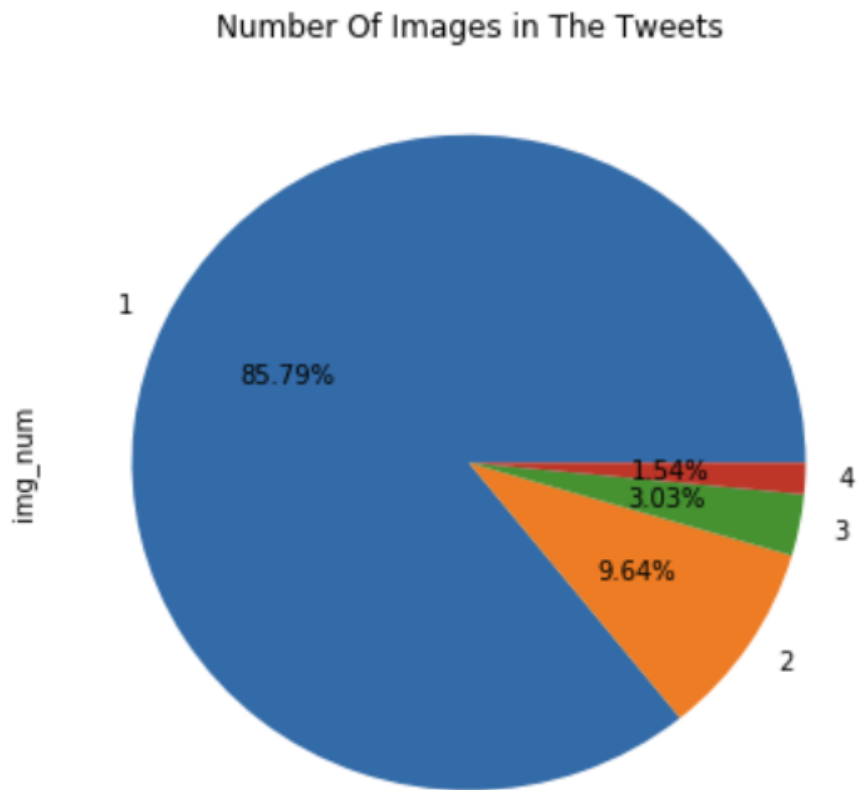
dog breed	count
golden retriever	154
labrador retriever	100
pembroke	94
chihuahua	90
pug	62

Question 2: What is the most source of tweets in the WeRateDogs Twitter archive?
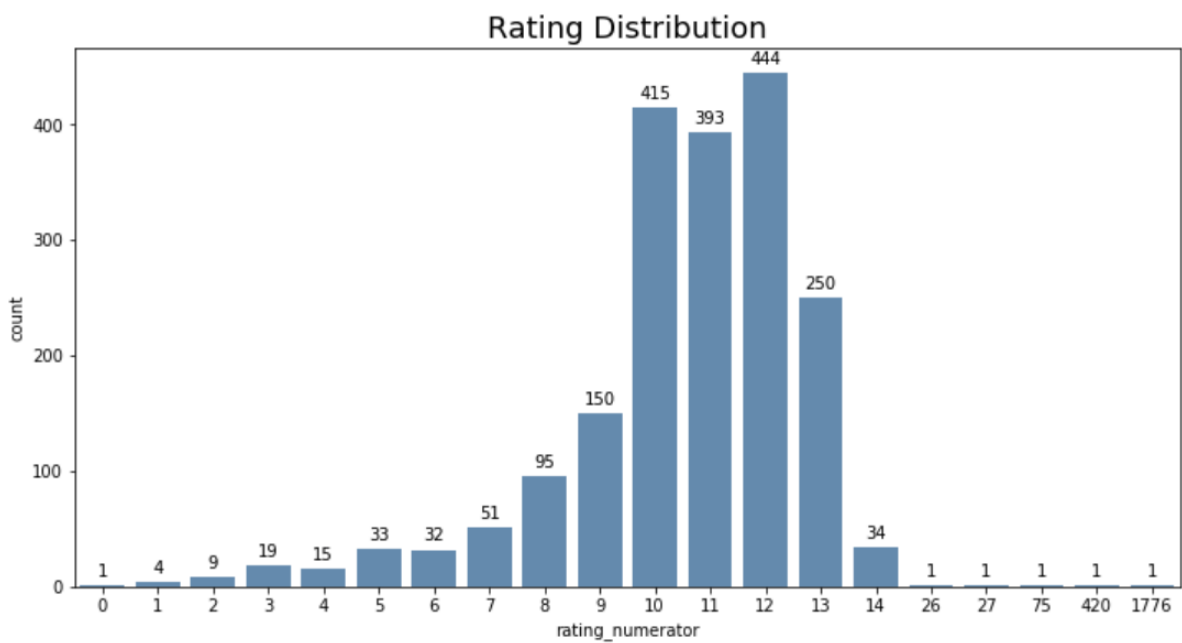
The dominant source of tweets in the WeRateDogs Twitter archive is Twitter for iPhone.

Question 3: What is the percentage of the number of images in tweets in our dataset?

Number Of Images in The Tweets

About 86% of the tweets in this archive have only one image, whereas only 1.5% have 4 images which is the maximum number of images in a tweet.

Question 4: What is the most rating that the dogs get in the dataset?

As the bar chart above shows the most rating that the dogs get in this archive is 12/10, Moreover there are some abnormal ratings like 26, 27, 75, 420, and 1776. but most of the ratings are from 10 to 12 all out of 10.