The WeRateDogs Project

Analysis, Visualization, and Insights from Twitter

By: Beltino Goncalves
Date: December 18, 2017

Introduction

If you haven't heard of them yet, well, I'll tell you about it. There are a twitter account @dog_rates, also known as WeRateDogs, with over 5.37 million followers that took over the internet early 2016 and it has received international media coverage.

WeRatedogs gained its popularity by rating people's dog images with a humorous comment about the dog. It was started by a college student, Matt Nelson with his friend when he set up a twitter pool from his personal account asking people to rate his dog. Via the positive response, it led him to create an empire for the internet's go-to dog rater of a scale from 1-10 range, with an exception of (13/10).

For this analysis, the data was composed of three pieces; WeRateDogs Twitter archive, tweet image predictions, and tweepy API json file. The WeRateDogs twitter archive was manually downloaded from this link: twitter_archive_enhanced.csv. This archive contains all @WeRateDogs 5000+ tweets from August of 2017. The tweet image prediction identified what breed of dog is present in each tweet according to a neural network. This file was hosted on the Udacity's servers and we downloaded it grammatically using python Requests library on the following link: here. The third piece of this process was the Tweedy API json file, which contains each tweet's retweeted count, favorite count and other additional that the team at Udacity found interesting. By using the tweet ID's in WeRateDogs twitter archive, a query was created to retrieve twitter API for each tweet's in JSON formatted data using Python's Tweepy library and then stored each tweet's entire set of a JSON data file called Tweet_json file.

Before diving into the visualization, and insights, I began by asking some basic question. What breed are most likely to be favorite and/or retweeted? Is there any animals other than dogs? What is the cutest doggo I could find?

You may be curious about who is this doggo image of on the left side. His name is Stephen, he holds the highest retweeted count and favorite count of all time. He had received a rating of 13/10 for his efforts. His owner tweeted, "This is Stephan. He just wants to help." He is the most popular dog in our dataset. He is classified as Chihuahua.

Sampling the dataset, and I found this pupper. Meet George, he is not your ordinary dog breed, but he has over 25K in favorites and 6K in retweets. His tweet said, "This is Georgie. He's very shy. Only puppears when called. Aggressively average at fetch. Unique front paws. Looks slippery. 10/10 would pet." This tweet was sent out at Mid-night on a Saturday.
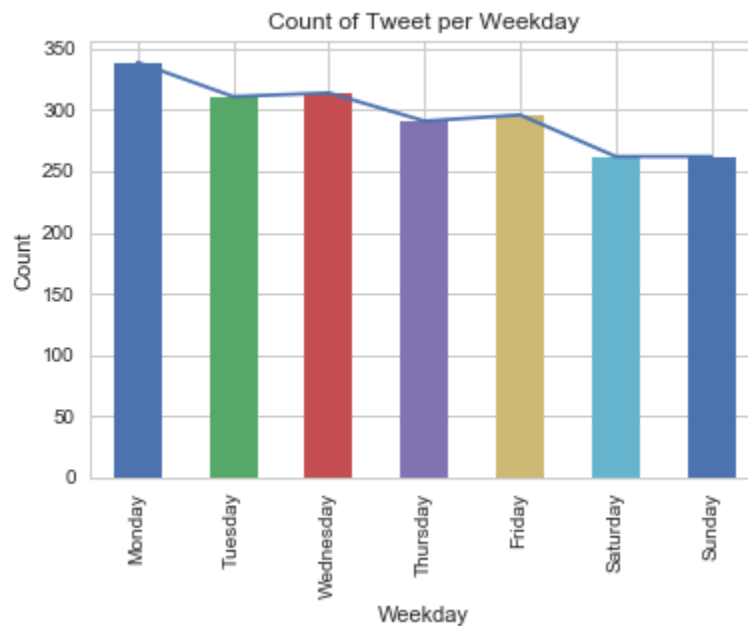

"This is Butter. She can have whatever she wants forever. 12/10 would hug softly." I would not disagree any further. She received over 30K on favorite count, also with 11K retweets. Sent out on a Wednesday.

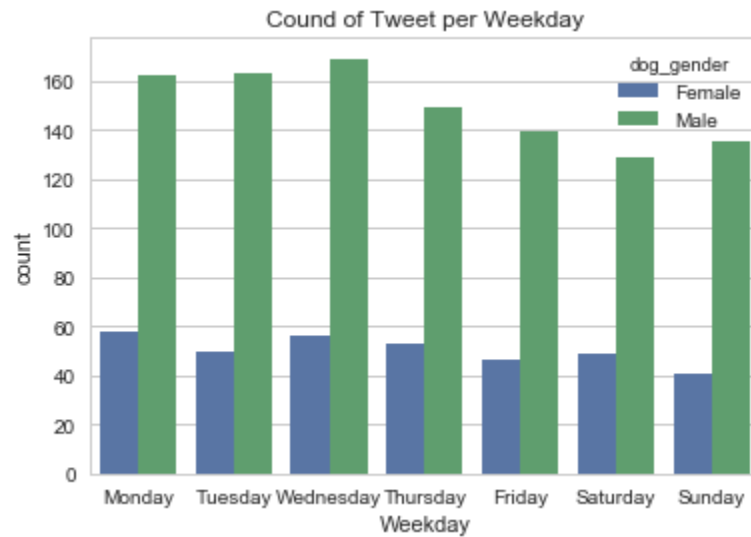# Visualization and Insights

Insight I

Count frequency of WeRateDogs Tweets per Weekday

Count of Tweet per Weekday

The Story

After spending a wonderful weekend with our dogs, Mondays are known to be the most crucial weekday, leaving us tweeting our dog's images and thinking of them while at work. Once the weekend hits again, as illustrated in the image above, we can now spend the time with our beloved dogs again on Saturday and Sunday, and post images @WeRatedogs.
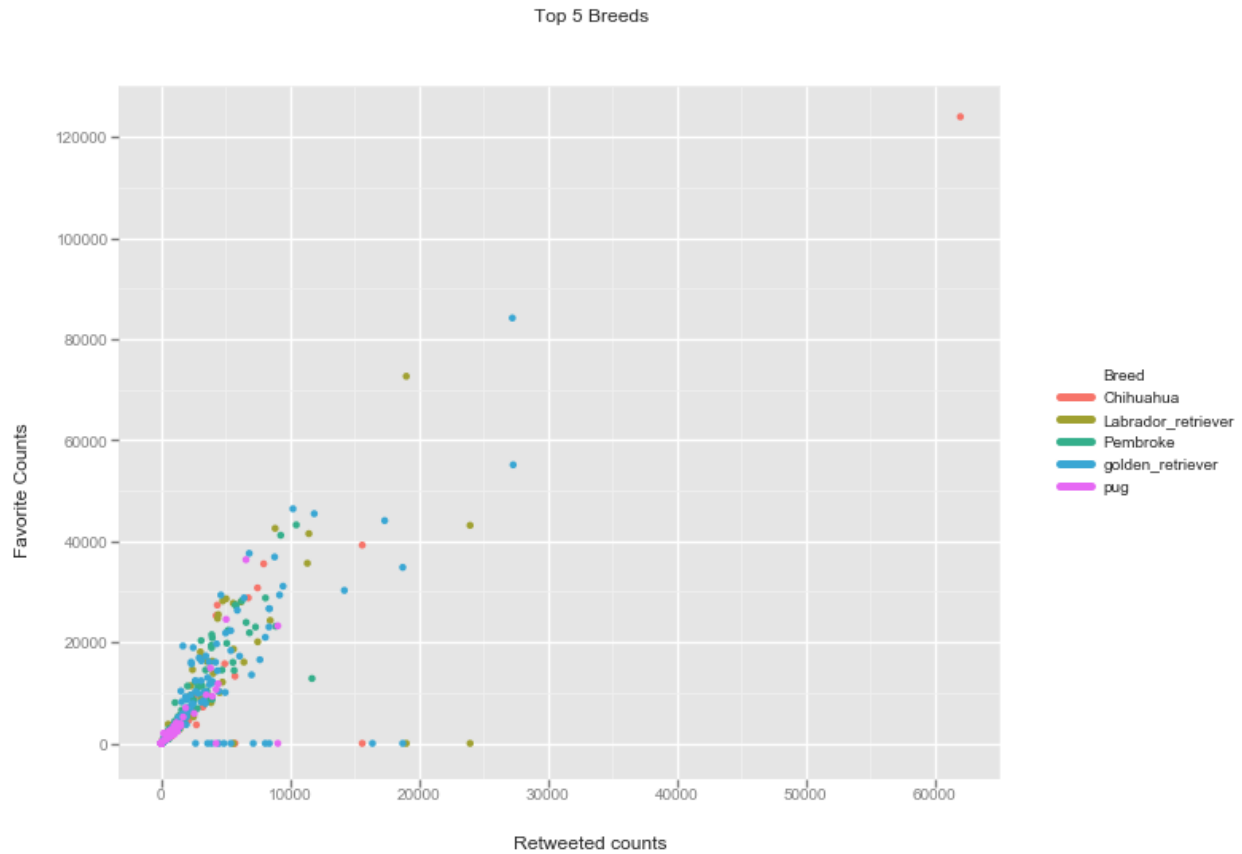
Gender Frequency count per Weekday

Cound of Tweet per Weekday

The Story

Throughout the weekday, male doggos were likely to be tweeted early in the weekday, leaving Wednesday being the highest count of tweets with decreasing male tweets for the rest for the weekday. For females, they were tweeted the most throughout the week on Monday and Wednesday, but the rate of change is between the weekdays are fairly small.

Insight II
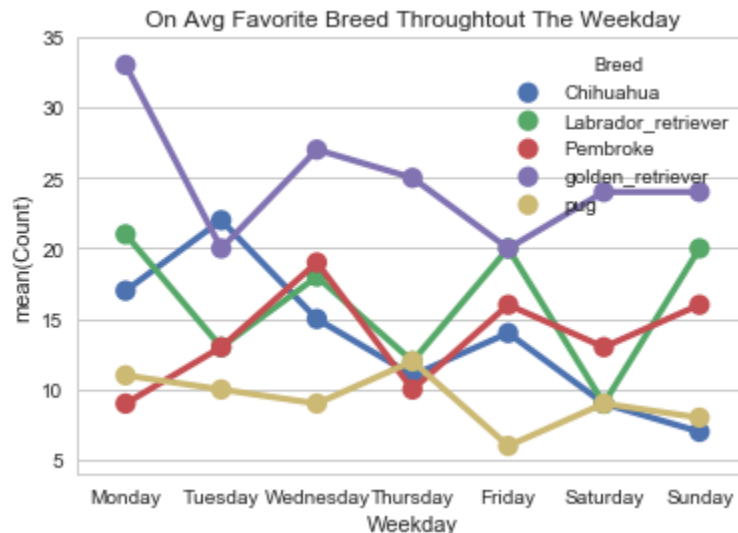
Top5 Breeds

Top 5 Breeds

The Story

Chihuahua's are by most the favorite and retweeted Breed, seconding golden retrievers, and with third Labrador retrievers. But take look at the Pug dogs, they are a very clustered group, compared with Pembroke and Labrador Retrievers which have more of a spread out based on favorite and retweeted scatterplot.

Insight III

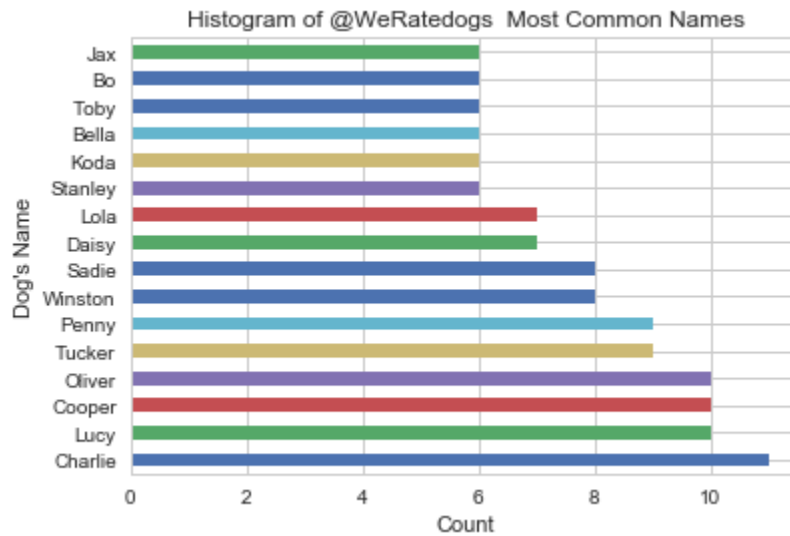***Average Breed tweeted per Weekday***

On average, the most tweeted @weratedogs breed throughout the week are Golden Retrievers. But on Tuesday the Chihuahuas takes the lead, as it could be referring to Taco Tuesdays. ☺ The Golden retrievers and Labrador retrievers meet on Fridays, to be the dogs most tweeted on average.

Breed	count	mean	std	min	25%	50%	75%	max
Chihuahua	95.0	10.521053	2.253444	3.00	10.0	11.0	12.0	14.0
Labrador_retriever	113.0	11.139381	1.441933	4.75	10.0	11.0	12.0	14.0
Pembroke	96.0	11.364583	1.870096	4.00	11.0	12.0	12.0	14.0
golden_retriever	173.0	11.550578	1.539915	0.00	11.0	12.0	12.0	14.0
pug	65.0	10.307692	1.802109	3.00	10.0	10.0	11.0	13.0

Again golden retrievers, on average gets the best ratings at 11.55. Making them @weratedogs the most beloved dog on the internet.
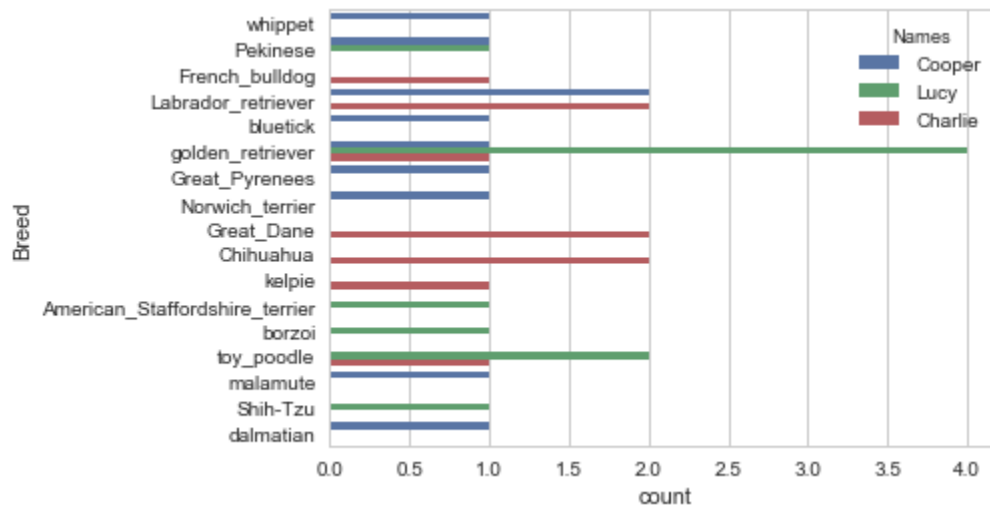
Insight IV

Dog's Common name

Histogram of @WeRatedogs Most Common Names

The Story

Charlie is the most common name for dogs, with Lucy, Cooper, and Oliver on Second.

Top 3 Common Dog Names based on Breed

Lucy is the most favorite dog name @WeRateDogs for golden retrievers. Leaving Charlie for chihuahuas with 2 counts, and Cooper for Labrador retriever for 2 counts.