

# Project 4 - Wrangle and Analyze WeRateDogs

Manuel Gajardo [megajardom@gmail.com](mailto:megajardom@gmail.com)

Mar 22, 2021

## Introduction

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. The account was started in 2015 by college student Matt Nelson, and has received international media attention both for its popularity and for the attention drawn to social media copyright law when it was suspended by Twitter for breaking these aforementioned laws.

As of December 2020, the Twitter account has nearly 9 million followers,[5] and Nelson sees 30,000 likes on a post as being viral.[1] His most popular post was of a dog marching in the 2017 Women's March, which was retweeted more than 50,000 times and favorited 134,000 times.[1]

The account's language has spawned an Internet language about "**doggos**" and "**puppers**".[6] A 2016 interaction with another Twitter user, when Nelson purposefully misnamed him "Brent" as is common in Weird Twitter, spawned the catchphrase "They're good dogs, Brent", which became one of the biggest memes of 2016.

<https://en.wikipedia.org/wiki/WeRateDogs>



# Project details

**Data wrangling WeRateDogs considered the following process**

- Gathering data
- Assessing data
- Cleaning data
- Storing, analyzing, and visualizing wrangled data
- Reporting on data wrangling efforts and data analyses and visualizations

I realized all the process mentioned above. I analyzed tweets using pandas and some libraries, methods and finally I used matplotlib to visualization.

The tools that I used were;

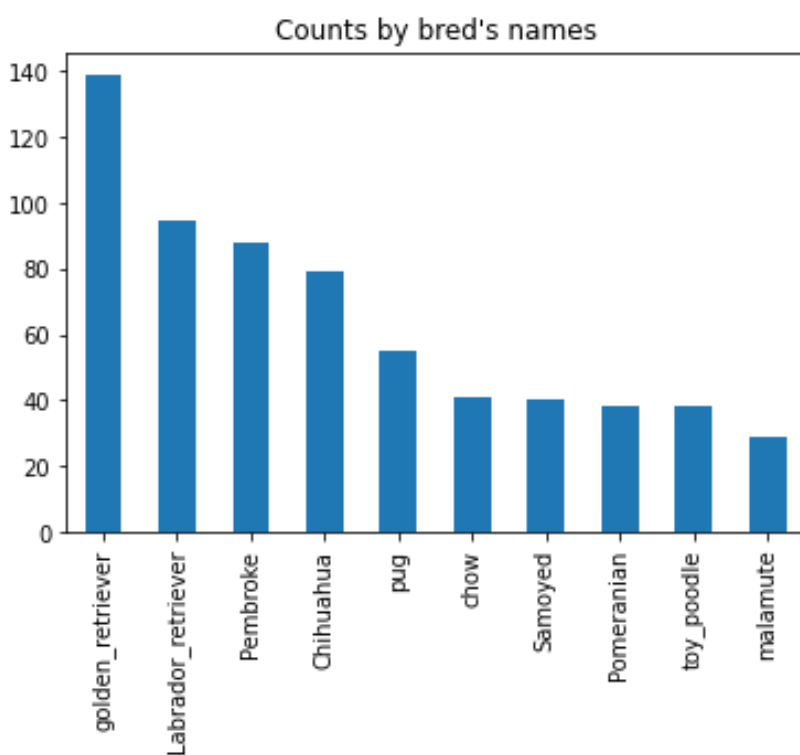
- Python
- Pandas library
- Request library
- Tweepy library
- Json library
- Matplotlib library
- Google Collab
- Anaconda Jupyter
- Tweeter API

# Insight and Visualization

## Insight #1

The most frequent breed of dog

golden_retriever	139
Labrador_retriever	95
Pembroke	88
Chihuahua	79
pug	55
chow	41
Samoyed	40
Pomeranian	38
toy_poodle	38
malamute	29

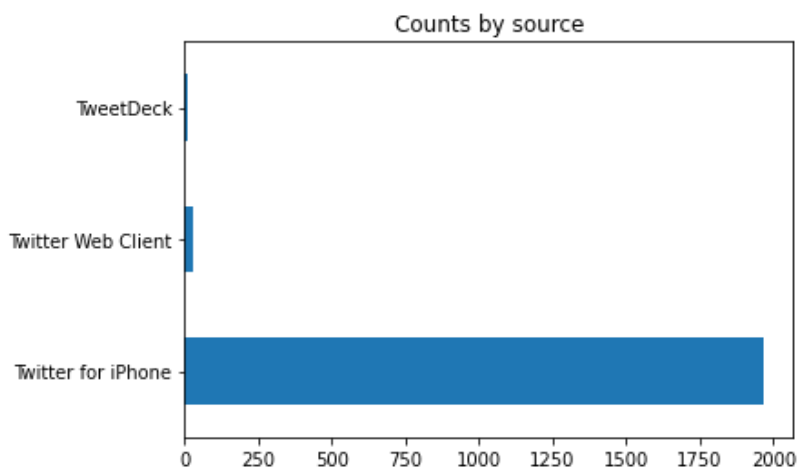


Golden retriever is the most frequent dog's breed

### Insight #2

The most frequent device from people is connecting

```
Twitter for iPhone    1968
Twitter Web Client    30
TweetDeck             11
Name: source, dtype: int64
```

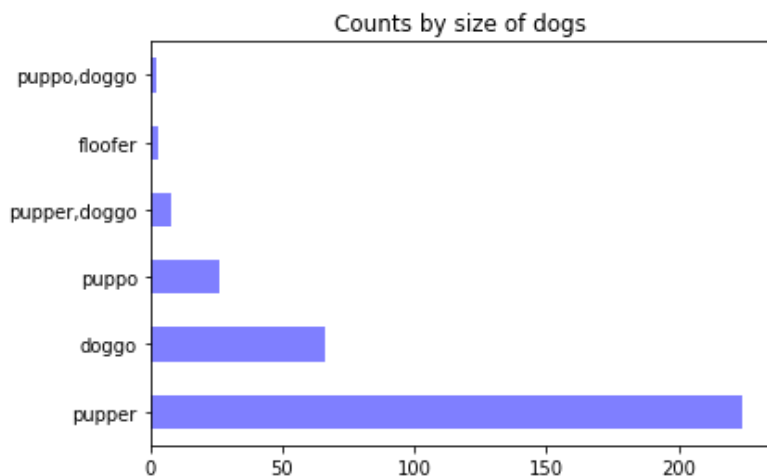


The graph show that the majority of people connecting is using iphone

### Insight #3

The most frequent stages of dogs are

```
pupper    224
doggo      66
puppo      26
pupper, doggo    8
floofer     3
puppo, doggo    2
```



The graph above show pupper is the most frequent stage of the dogs

#### Insight #4

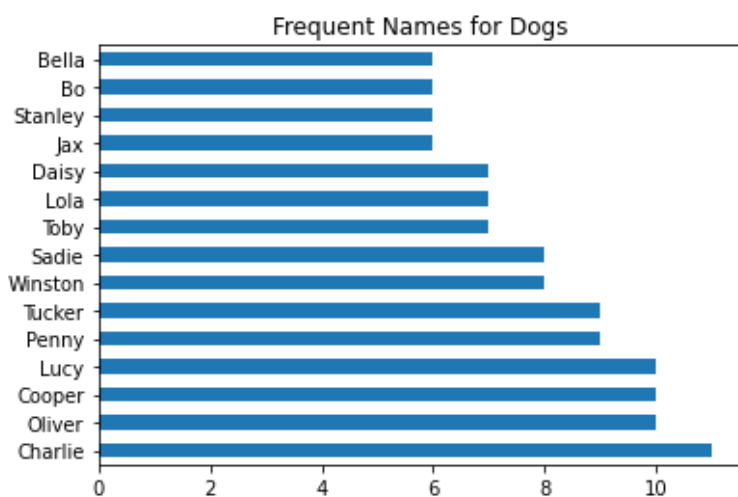
The less frequent predicted dog's bred

canoe	1
terrapin	1
cheeseburger	1
cliff	1
swab	1
school_bus	1
ibex	1
pitcher	1
beaver	1
hummingbird	1

#### Insight #5

The most frequent dog's name

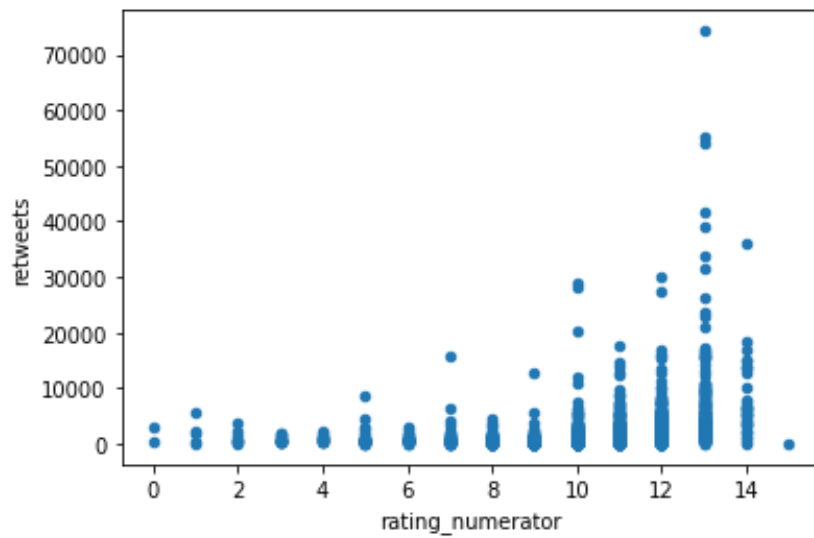
Charlie	11
Oliver	10
Cooper	10
Lucy	10
Penny	9
Tucker	9
Winston	8
Sadie	8
Toby	7
Lola	7
Daisy	7
Jax	6
Stanley	6
Bo	6
Bella	6



The graph above show that CHARLIE is the name more frequent

### Insight #6

Rating v/s Retweet relationship



The graph shows a relation between rating and retweets, higher rating -> higher retweets