

# Data Analysis and Visualization Report: WeRateDogs

By: Allison Darrington



## Introduction

The image shows three tweets from the WeRateDogs Twitter account (@dog\_rates) in a vertical stack. The first tweet is from user @brant, dated 12 Sep 2016, stating "@dog\_rates your rating system sucks. Just change your name to "CuteDogs"." The second tweet is from the WeRateDogs account (@dog\_rates), dated 12 Sep 2016, responding to @brant with "Why are you so mad Bront". The third tweet is from user @brant, dated 12 Sep 2016, replying to @dog\_rates with "well you give every dog 11s and 12s. It doesn't even make any sense." Each tweet includes engagement metrics: likes, retweets, and replies.

The image shows a single tweet from the WeRateDogs Twitter account (@dog\_rates). The tweet text is "@brant they're good dogs Brent". Below the tweet are engagement metrics: RETWEETS (6,616), LIKES (16,335), and a row of small thumbnail images. At the bottom, it shows the timestamp "4:05 PM - 12 Sep 2016" and additional engagement metrics: 88 replies, 6.6K retweets, and 16K likes. A "Following" button is visible next to the account name.

Oh, Brunt. All puppers are heckin' good, therefore they all deserve a rating of at least 10/10. This is the motto of [WeRateDogs](#) Twitter account, which uses a unique rating system and "doggo language" to share and celebrate dogs from all over the world. Their cute and quirky ratings and captions capture the hearts of nearly 8.9 Million followers.

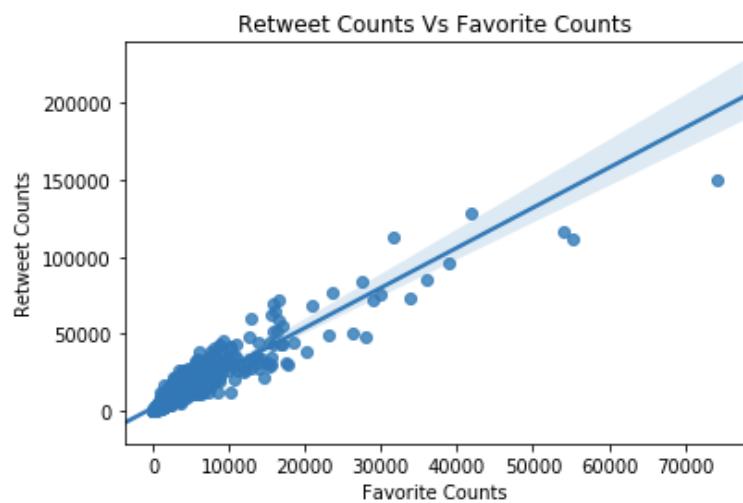
Some questions you might ask when scanning through WeRateDog's tweets: Who is the most popular dog? What types of breeds are seen the most? Which dog has the highest rating? Through a data analysis, all of these questions

were answered, plus more. Tweet information gathered runs from November of 2015 to August of 2017.

## Analysis and Visualizations

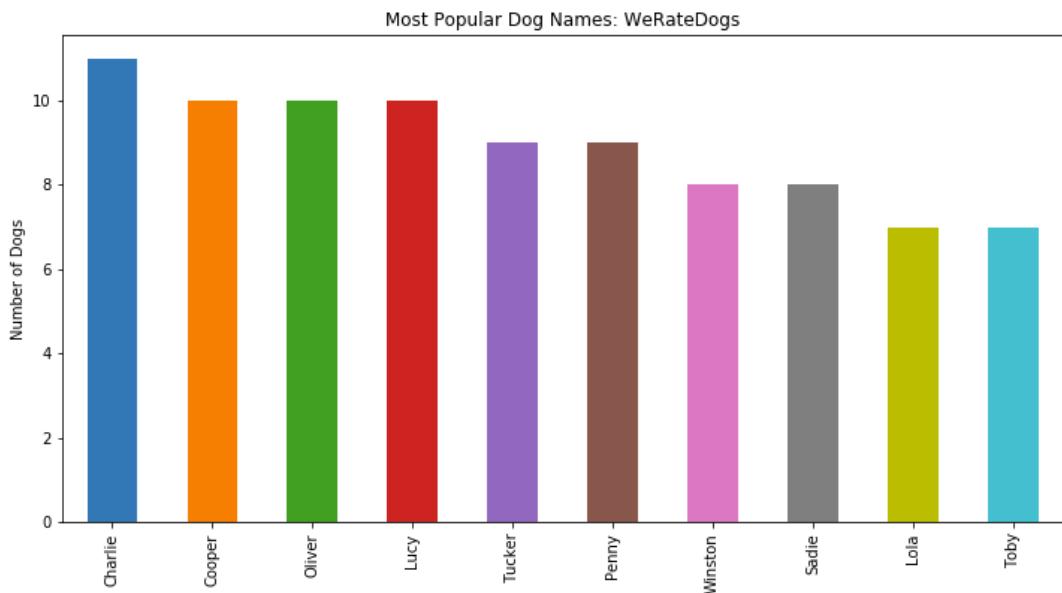
All the data wrangling, analysis and visualizations was done using Python 3. See Jupyter Notebook wrangle\_act.ipynb for the coding used to gather this information.

### What is the correlation between retweets and favorites?



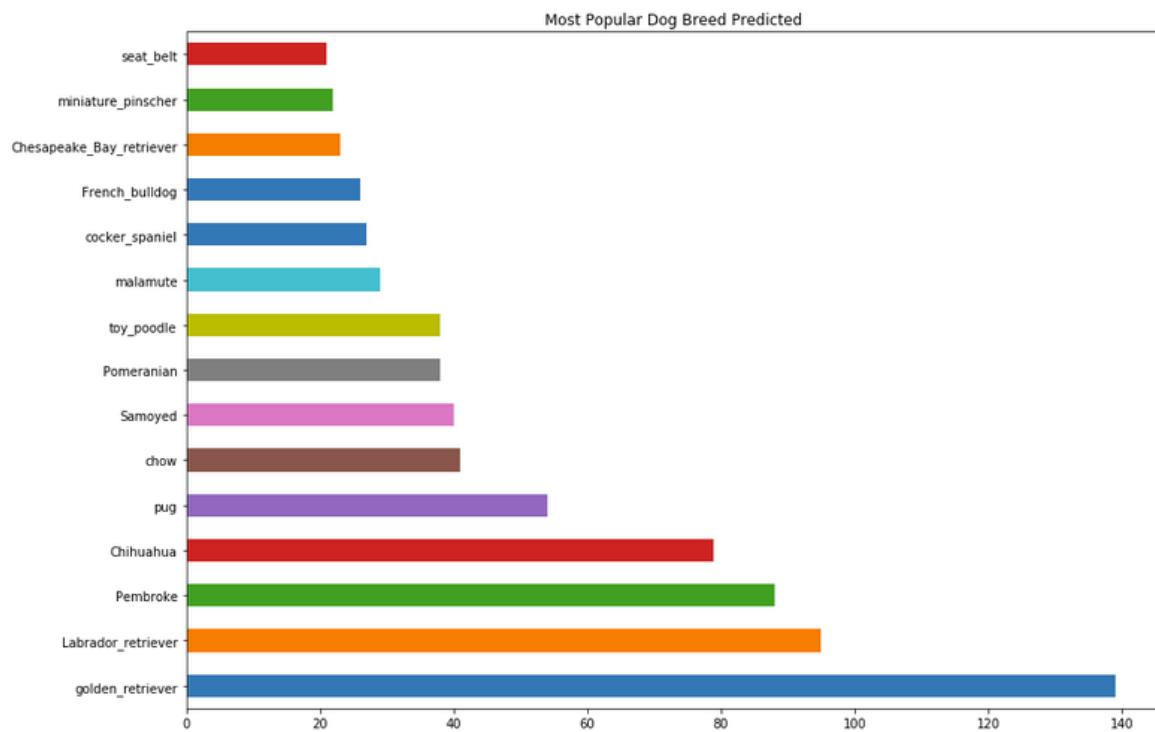
There is a strong positive relationship between retweet counts and favorite counts.

### Most Popular Dog Names?



Charlie is the most popular dog name, followed by Cooper.

## Most Popular Breeds?



An algorithm was used to match breeds with their images. Golden Retriever was predicted the most.



This is Stuart, the first image in the database recognized as a Golden Retriever.  
The algorithm predicted this with 95% confidence. Looks like it was right!

## Dog with the most favorites?

The most favorited tweet is:

[Here's a doggo realizing you can stand in a pool. 13/10 enlightened af \(vid by Tina Conrad\)](#)



### Dog with the most retweets?

The most retweeted tweet was the same one who had the most favorites, pool doggo! This is evidence of the strong relationship between retweets and favorites.

### Dog with the highest rating?

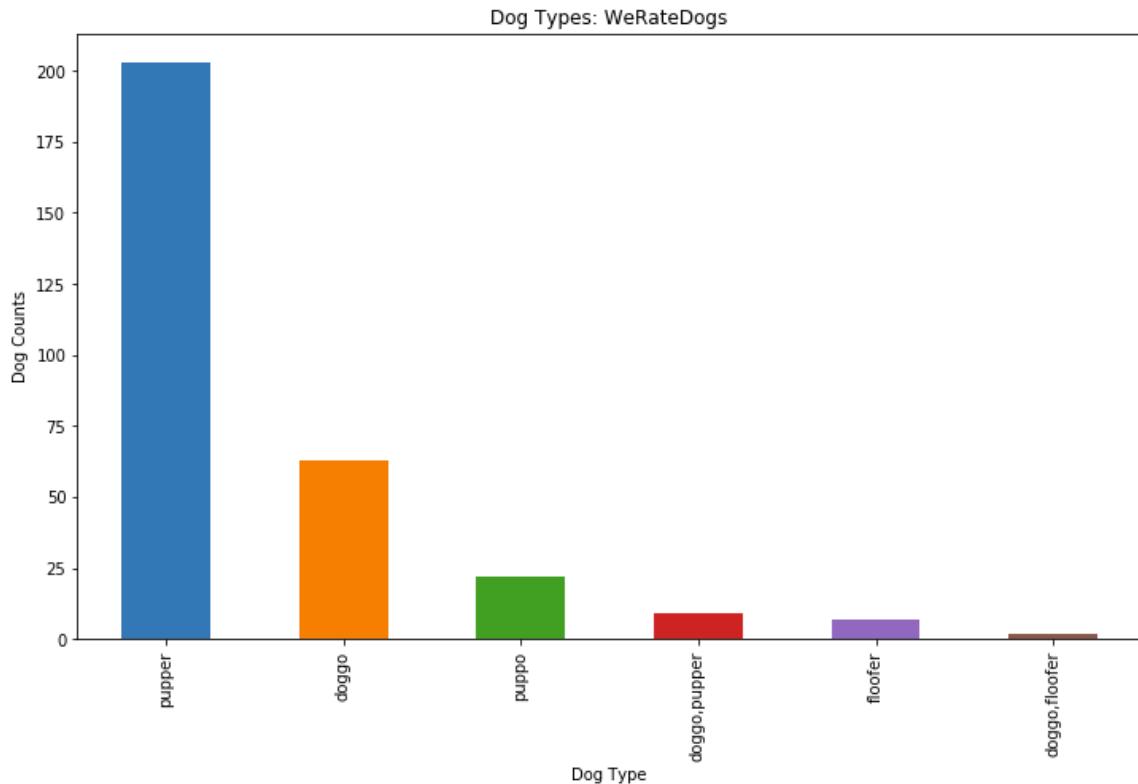
This dog with the highest rating is:

[This is Atticus. He's quite simply America of. 1776/10](#)



1776/10 is well deserved!

Doggo, Floofer, Pupper or Puppo? Which type is represented the most?



This concludes the questions answered in this data analysis. To finish off, are 5 more images of a random sample from the dataset. Because we all want to see more dogs!



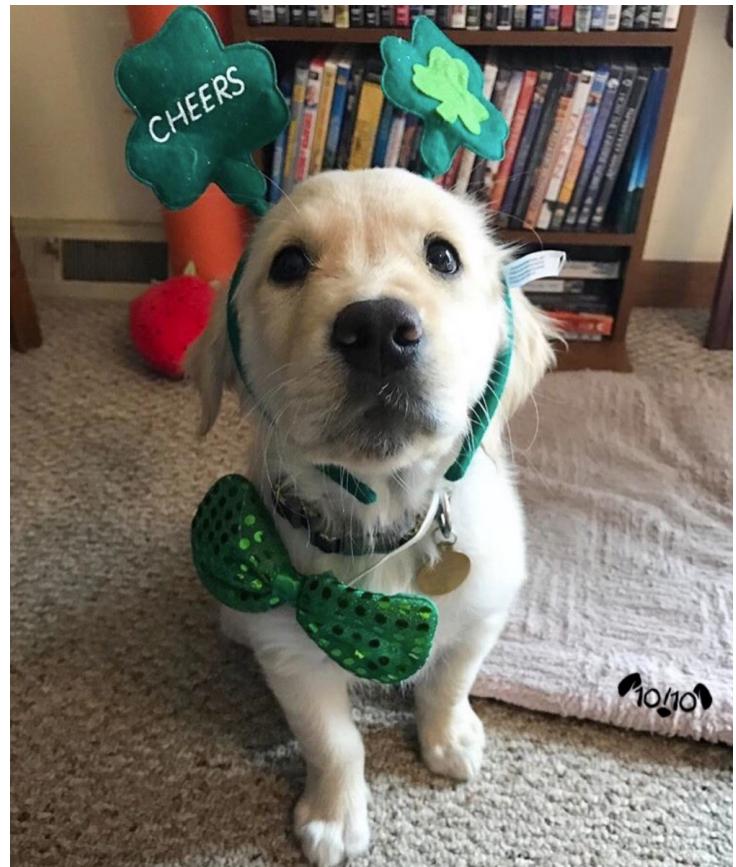
This is Chef. Chef loves everyone and wants everyone to love each other. 11/10

This is Bella. She's a Genghis Flopped Canuck. Stuck in trash can. 9/10 not to happy about it



\*screams for a little bit and then crumples to the floor shaking\* 12/10

This is Charlie. He's wishing you  
a very fun and safe St.  
Pawtrick's Day. 13/10 festive of



This little pupper just arrived.  
11/10 would snug

## Conclusion

With proper data wrangling, assessing, cleaning and visualizing, we can learn unique insights from this twitter account. Without proper cleaning, many of the insights gathered in this document would not have been as apparent.

An example of what could have been missed without proper data gathering: Breed Prediction table. This was an additional dataset that brought a lot of value by showing the most represented breeds.

Additionally, an example of what could have been missed without proper data cleaning: Some of the numerator and denominator values were incorrect. These are all heckin' good pups, and they all deserve their true ratings.

This project emphasizes how important it is to gather your data, assess and clean it properly to get the most accurate results.