

Analyzing, and Visualizing Data

By: Kim Gahn

I have learned much throughout my time in the Udacity Data Analysis program, but some of the most powerful tools have come from the Data Wrangling course. The capstone project has given me the ability to query data from a popular social media platform in order to provide insightful feedback and creative visualizations.

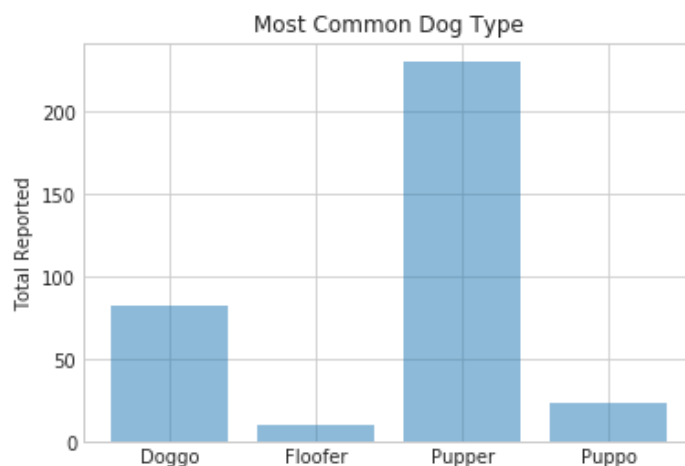
This project focused on the Twitter data for the user WeRateDogs, that rates people's dogs with a humorous comment about the dog. WeRateDogs has over 4 million followers and has received international media coverage.

One aspect that I thought was unusual was descriptive categories in which many dogs are placed. These include:

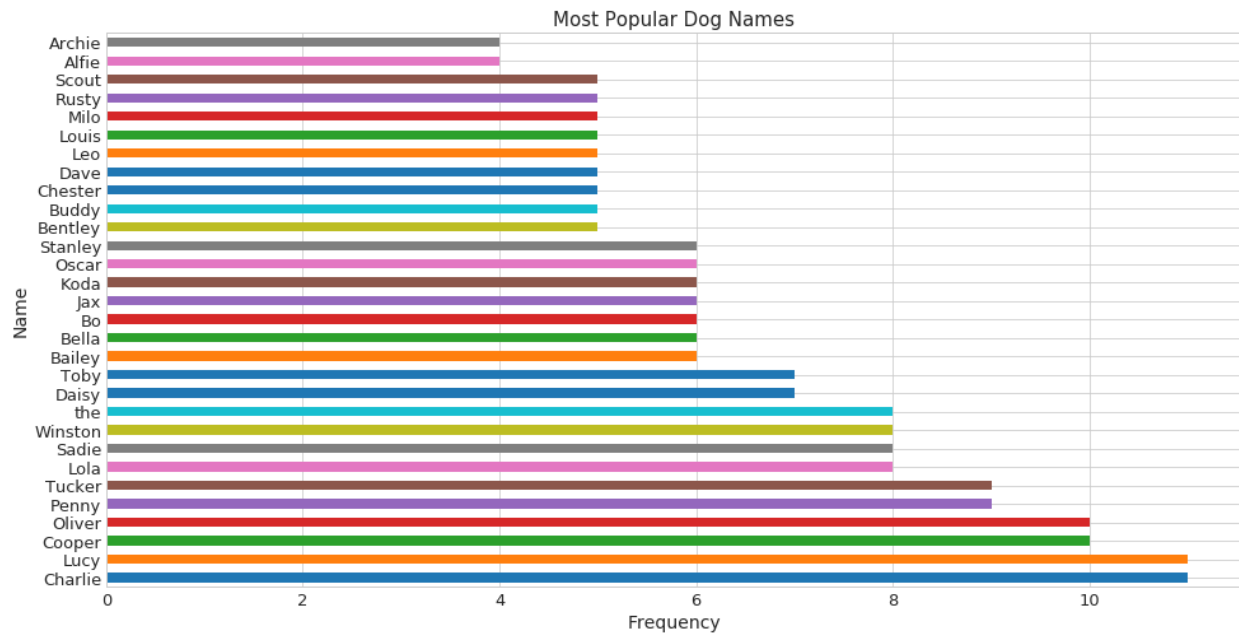
- Doggo – The word “doggo” evokes a dependable, tail wagging friend that will stay closer than your best friend. (Note: All dogs can be doggos. But not all doggos are puppies, woofers or floofers.) ¹
- Floofer – Large hairy or fluffy dogs. ¹
- Pupper – A small dog or puppy. ¹
- Puppo – An adolescent dog. ¹

1 - <https://www.rover.com/blog/uk/do-you-speak-doggolingo-dog-internet-jargon-explained/>

So which of these types is most noted in our data? The puppies have it!

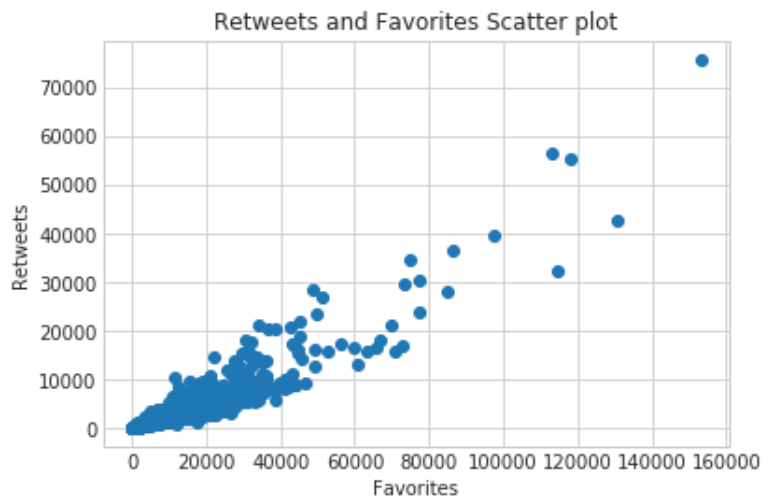


Next I wanted to know what the most popular dog names were. Could it be Fido or Snoopy, or something else that we normally only associate with a dog?



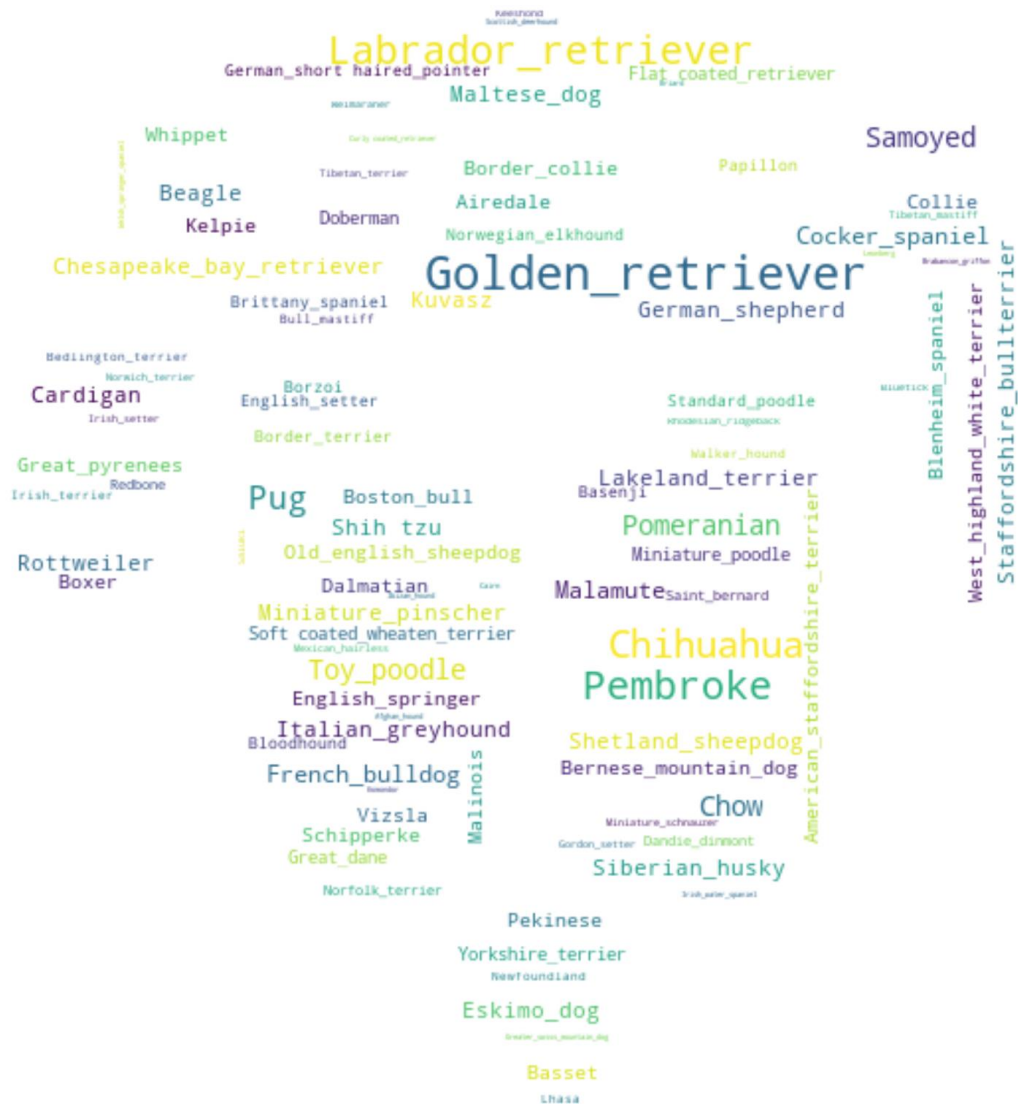
NO! The list of the top 30 names seem to be names that any human could use.

Then I wanted to know if a person marks a tweet as their favorite, what is the likelihood that they would retweet it?



The strong almost straight-line appearance in the above plot shows that it is very likely.

Finally, I wanted to know a bit about what dog breeds appear in our data.



It looks like some of the most popular breeds are the golden retriever, lab, and Chihuahua. That makes sense, I have a golden of my own!

There is so much power in being able to take data from multiple sources and bend it to your will, and I can't wait to include these techniques in future projects.