

Data Analysis

Once the data has been cleaned and wrangled into something that can be worked with, it can be used to garner insights.

The three insights that I wanted to discover were:

1. I wanted to see the distribution of retweet count by dog breed and see the top 3 and bottom 3 breeds. The insight that I observed was the most popular 3 breeds were the Briard, Irish Setter, and the Irish Water Spaniel while the bottom 3 retweets were the Tibetan Terrier, the Redbone, and the Black and Tan Coonhound.
2. The second insight that I wanted to observe is if the previous insight would match up with the favorites. Upon closer inspection, the top 3 favorited breeds changed up some but the bottom 3 breeds were consistent, The Tibetan Terrier, The Redbone, and The Black and Tan Coonhound.
3. The last insight I wanted to examine was the most popular dog names. While looking at this data, I noticed multiple entries where the dog's name was "a" or "the". This is most likely a typo or a placeholder for the dog's name. I displayed the records to double check and then I dropped them.

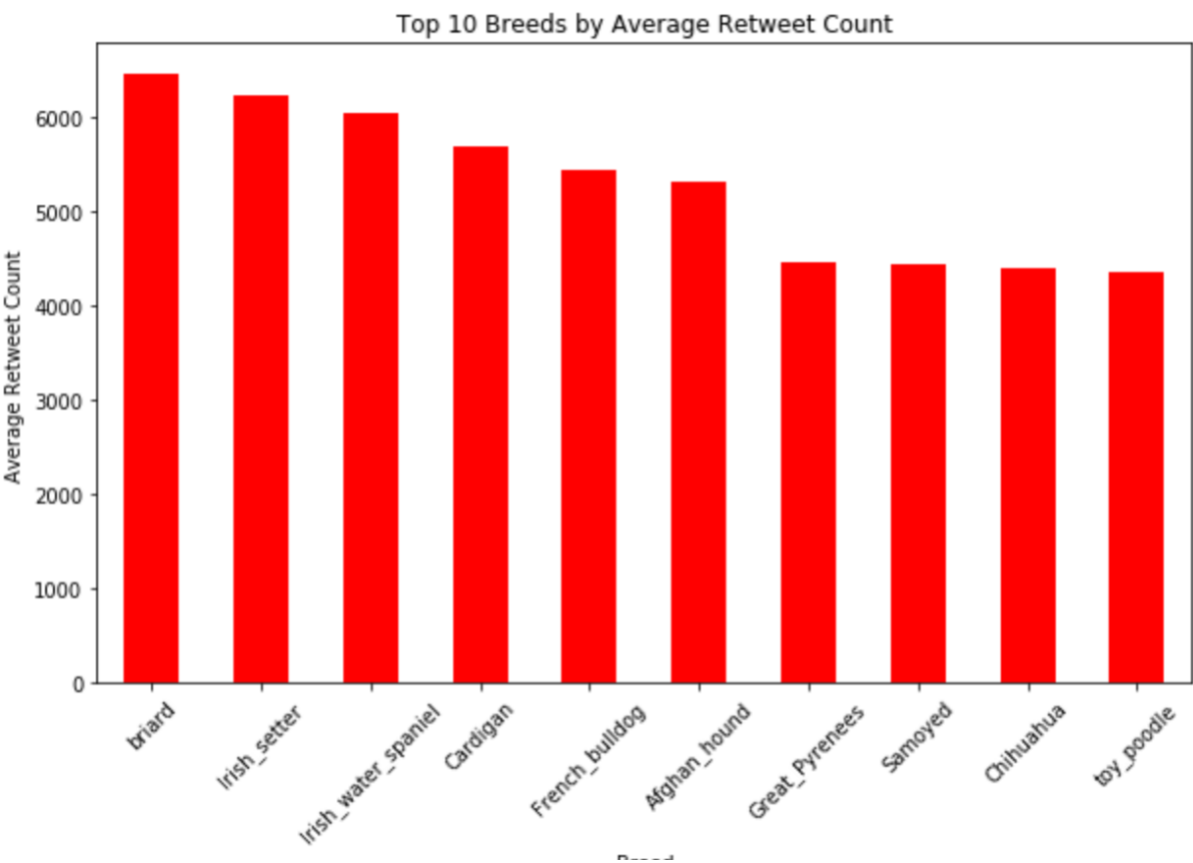
Most and Least Popular dog breeds by Retweet Count

The most popular dog breed by retweet count was the Briard with an average of 6470.

	count	mean	std	min	25%	50%	75%	max
breed								
briard	1.0	6470.0	NaN	6470.0	6470.0	6470.0	6470.0	6470.0
Irish_setter	1.0	6238.0	NaN	6238.0	6238.0	6238.0	6238.0	6238.0
Irish_water_spaniel	1.0	6039.0	NaN	6039.0	6039.0	6039.0	6039.0	6039.0

The least popular dog breed by retweet count was the Black and Tan Coonhound with only 44.

	count	mean	std	min	25%	50%	75%	max
breed								
Tibetan_terrier	1.0	66.0	NaN	66.0	66.0	66.0	66.0	66.0
redbone	1.0	48.0	NaN	48.0	48.0	48.0	48.0	48.0
black-and-tan_coonhound	1.0	44.0	NaN	44.0	44.0	44.0	44.0	44.0



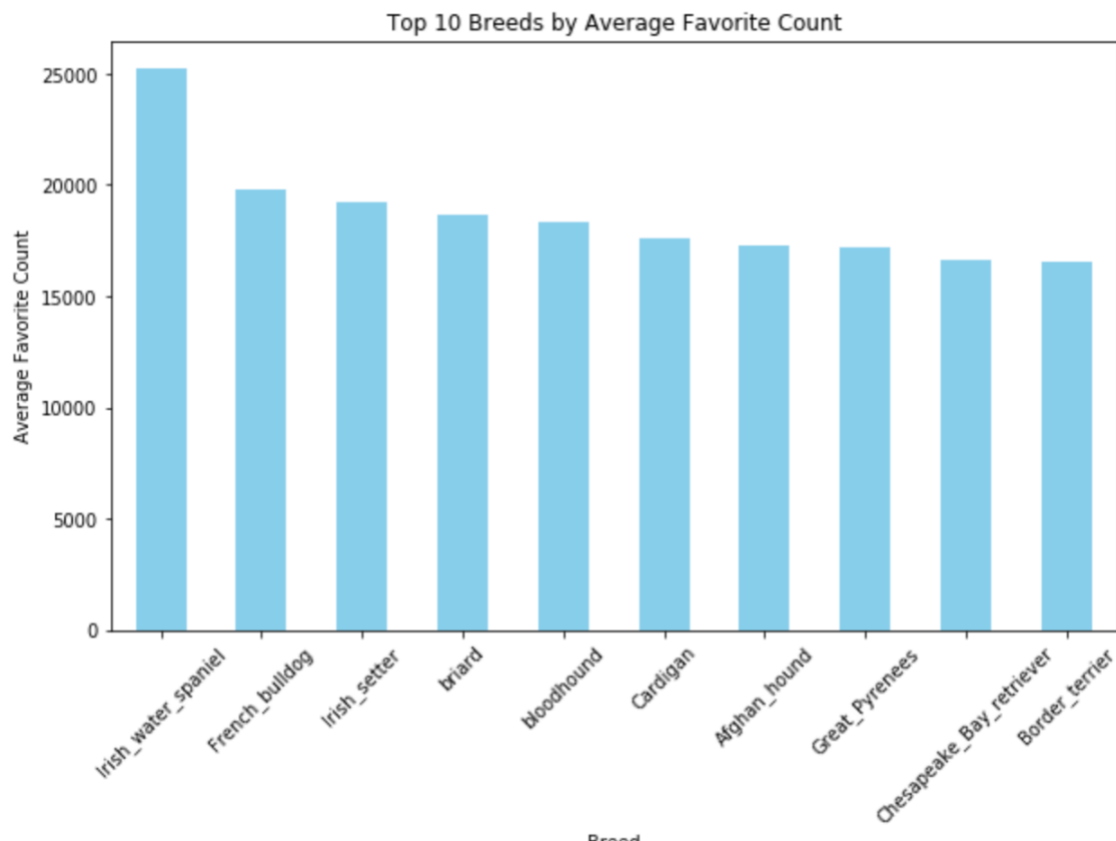
Most and Least Popular dog breeds by Favorite Count

I expected to see similar results when looking by Favorite Counts. Surprisingly the average most ‘favorited’ dog breed was the Irish Water Spaniel.

```
breed
Irish_water_spaniel    25225.000000
French_bulldog        19811.947368
Irish_setter           19250.000000
Name: favorite_count, dtype: float64
```

The least popular dog breed by favorite was more in line with what I was expecting with the Black and Tan Coonhound having an average of 113 'favorites'.

```
breed
Tibetan_terrier        203.0
redbone                132.0
black-and-tan_coonhound 113.0
Name: favorite_count, dtype: float64
```



Most popular dog names

When I wanted to determine the most popular dog names I displayed the top ten most popular names. This displayed some additional cleaning that I had to do. Some records did not have a name and only had "a" or "the" in the name column. I proceeded to drop these records and came with the list that shows Oliver was the most popular with Cooper and Penny tied for second..

Oliver	7
Cooper	6
Penny	6
Lucy	5
Stanley	4
Winston	4
Charlie	4
Koda	4
Louis	3
Winnie	3

Name: name, dtype: int6

Further visualizations showed that these **10** names accounted for **46%** of the dog names analyzed.

