



Wrangle and Analyze Data

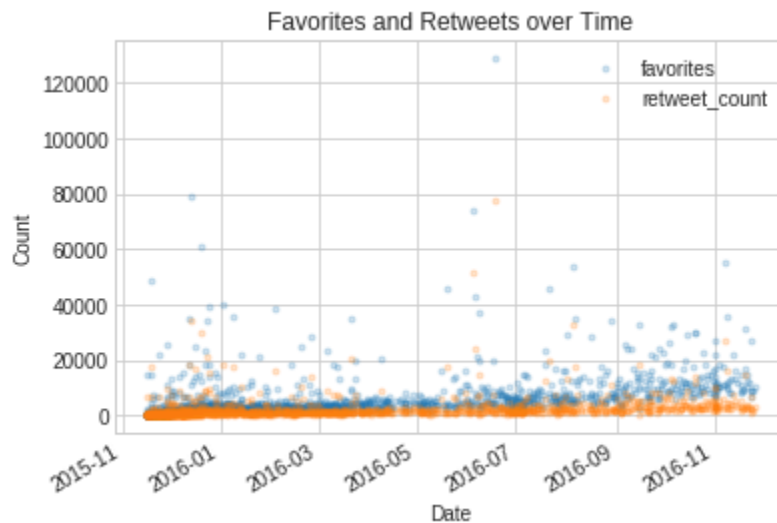
ANALYSIS

Yogeshwar Murugan | Udacity - DAND | 24-05-18

INTRODUCTION:

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators however are almost always greater than 10 (11/10, 12/10, 13/10, etc.). In theory the rates should be 1 to 10. However, WeRateDogs disagree, they admit almost all dogs deserve a 10 and sometimes more than that. WeRateDogs has over 6 million followers and has received international media coverage. One of those coverage were about the quote “they’re good dogs Brent”, this was an exchange in which WeRateDogs shut down a person having an issue with its rating system in humorous ways. So what does this rating do? Which dog type (breed) is most common dog in the tweet dataset? What dog type has the highest average rating? These questions and more are answered in the following insights:

Insight 1:



There seems to be an increase in the page's overall 'favorites' and 'retweet' counts over the course of 1 year. This clearly depicts that the page has grown popular over this duration. More 'favorites' and 'retweets' also means that people started liking the page and were visiting and sharing it much often, contributing to its increase in popularity.

Insight 2:



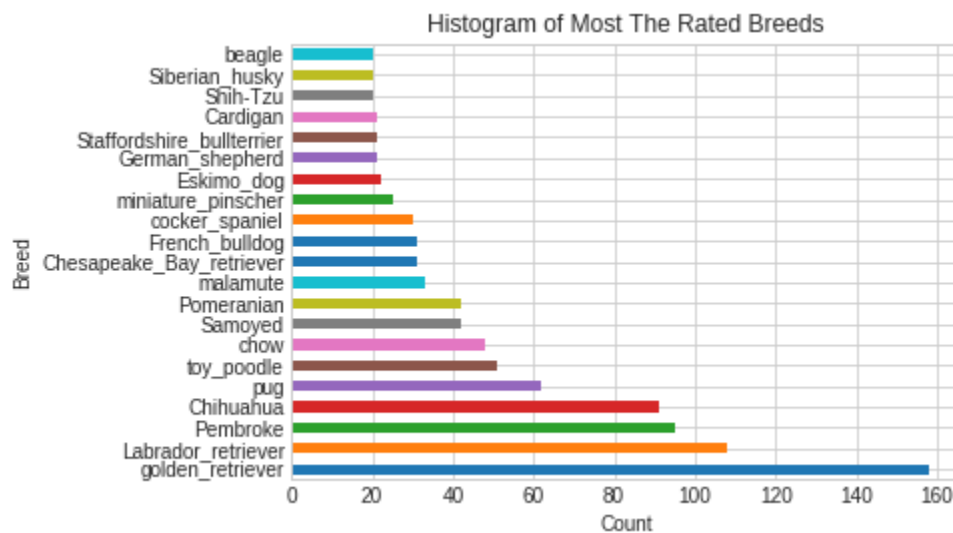
The ratings for the dogs, irrespective of the breeds, seems to increase over time. More number of High ratings were given to the dogs as time progresses. This could mean either the dogs being rated were 'Good Dogs', or either the page started giving off good numbers and funny comments for improving their popularity.

Insight 3:

	count	mean	std	min	25%	50%	75%	max
breed								
Cardigan	21.0	11.142857	1.590148	7.0	10.00	11.0	12.0	13.0
Chesapeake_Bay_retriever	31.0	10.838710	1.439833	8.0	10.00	10.0	12.0	13.0
Chihuahua	91.0	10.516484	2.083601	3.0	10.00	11.0	12.0	14.0
Eskimo_dog	22.0	11.409091	1.402688	9.0	10.00	12.0	12.0	14.0
French_bulldog	31.0	11.193548	1.796652	8.0	10.00	12.0	12.0	14.0
German_shepherd	21.0	11.000000	1.449138	8.0	10.00	11.0	12.0	13.0
Labrador_retriever	108.0	13.638889	16.791998	7.0	10.00	11.0	12.0	165.0
Pembroke	95.0	11.410526	1.716594	4.0	11.00	12.0	12.0	14.0
Pomeranian	42.0	12.476190	9.978607	6.0	10.00	11.0	12.0	75.0
Samoyed	42.0	11.690476	1.352290	7.0	11.00	12.0	13.0	14.0
Shih-Tzu	20.0	10.350000	1.308877	8.0	9.75	10.5	11.0	13.0
Siberian_husky	20.0	13.250000	8.723079	9.0	10.75	11.5	12.0	50.0
Staffordshire_bulldog	21.0	10.761905	1.374946	8.0	10.00	11.0	12.0	13.0
beagle	20.0	10.150000	1.531253	6.0	9.75	10.0	11.0	13.0
chow	48.0	14.166667	19.185914	7.0	11.00	12.0	12.0	144.0
cocker_spaniel	30.0	11.333333	1.184187	9.0	10.25	11.5	12.0	13.0
golden_retriever	158.0	13.000000	12.623404	5.0	11.00	12.0	12.0	143.0
malamute	33.0	10.878788	1.218544	8.0	10.00	11.0	12.0	13.0
miniature_pinscher	25.0	10.000000	2.565801	2.0	9.00	11.0	12.0	12.0
pug	62.0	10.241935	1.816910	3.0	9.25	10.0	11.0	13.0
toy_poodle	51.0	11.039216	1.264291	7.0	10.00	11.0	12.0	13.0

The 'Chow' breed of dogs seem to have the highest mean Rating among all the breeds followed by 'Labrador_retriever' and by 'Siberian_husky'. This insight can be confirmed true as it is evident to everybody that the above mentioned three breeds of dogs are very friendly. On the other side of the scale, 'miniature_pinscher' has received the worst ratings (2) of all the breeds. It also holds a name for getting the lowest rating of all time.

Visualization 1:



This histogram gives us a insight on the population of dogs that were rated. It is lead by 'golden_retriever' with almost 160 ratings. This can be insight can be confirmed, because Golden retrievers are the most common domestic dog pets. 'Beagle' has the lowest count in number of all the dogs that were rated.