

## Analyzing WeRateDogs Twitter Archive

September 13<sup>th</sup>, 2022

The following questions about the dataset were raised from the variables available in the dataset;

1. What breed has more engagement?
2. Does the time of posting affect the level of interaction with each tweet?
3. Is a dog stage associated with more ratings

### What breed is associated with the most engagements?

Are some breeds more attractive to WeRateDogs followers? To answer this, we should explore each breed's favorite and retweet count. The counts are illustrated below;

	tweet_id	rating_numerator	rating_denominator	favorites	retweets	img_num	confidence_lvl
breed							
saluki	8.587015e+17	13.000000	10.0	30932.000000	6508.000000	1.000000	0.535929
irish_water_spaniel	8.453069e+17	12.000000	10.0	25225.000000	6039.000000	1.000000	0.567475
french_bulldog	7.772229e+17	11.227273	10.0	18099.363636	4933.954545	1.136364	0.847845
afghan_hound	8.413147e+17	13.000000	10.0	17305.000000	5312.000000	1.000000	0.903712
eskimo_dog	7.716185e+17	11.888889	10.0	17148.000000	7826.555556	1.333333	0.576585
...	...	...	...	...	...	...	...
english_springer	6.736288e+17	10.666667	10.0	804.000000	264.000000	1.000000	0.853234
miniature_schnauzer	6.686314e+17	5.000000	10.0	763.000000	349.000000	1.000000	0.904549
redbone	6.685126e+17	9.500000	10.0	653.500000	182.500000	1.000000	0.686523
giant_schnauzer	6.671713e+17	10.000000	10.0	235.000000	97.000000	1.000000	0.841265
tibetan_terrier	6.674554e+17	7.000000	10.0	203.000000	66.000000	1.000000	0.676376

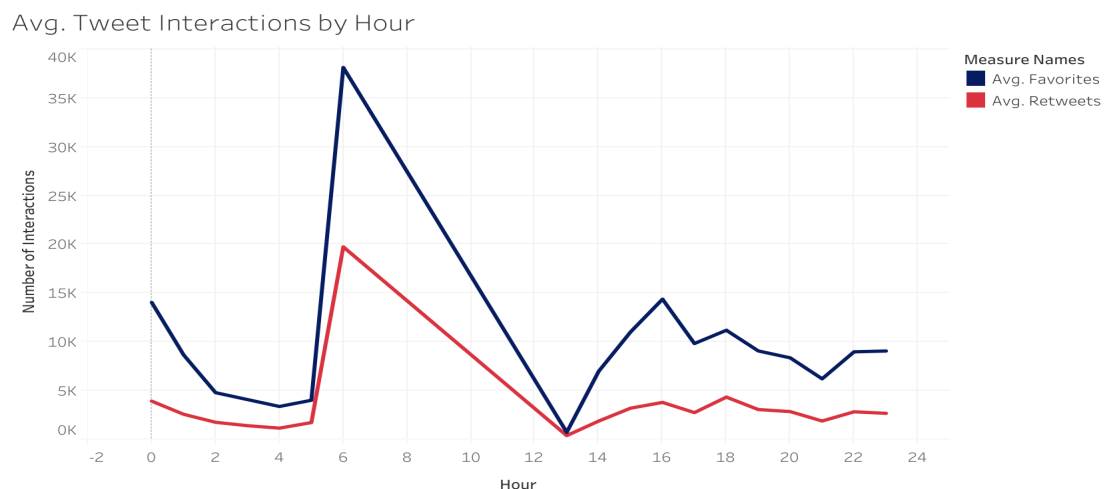
100 rows × 7 columns

**Figure 1:** Average favorites and retweets for each breed

Here, we see a mean summary of each predicted breed with a confidence level above 50%. The image shows that Salukis have the most engagements with 30,932 favorites, and 6,508 retweets at an average.

### Does the time of posting affect the level of interaction?

Intuition tells us that the time tweets are posted can determine the level of interactions from followers. We can measure interactions by the favorite and retweet counts of each tweet. To arrive at a conclusion, the hour each tweet was posted has been extracted from the timestamp, and grouped by each hour of the day. A mean representation of favorites and retweets by each hour are illustrated in the chart below:



**Figure 2:** Line plot of average favorites and retweets by each hour of the day

The line plot above tells us that on average, within the hour of 6am, there's a spike in engagements, and it starts to decline through midday, hitting its lowest favorite and retweet counts at the 1pm mark.

### Is a dog stage associated with more ratings?

We would like to get an idea of how much ratings a dog can get from WeRateDogs. The stage variable may provide insights to this. Does a dog have a chance of getting high ratings if it's a doggo, pupper, puppo or floofer?

Figure 3, is a result of grouping the data into stages. We can observe that floofer and puppo have the more ratings with a rating numerator of 12.

	tweet_id	rating_numerator	rating_denominator	favorites	retweets	img_num	confidence_lvl	hour
stage								
doggo	8.007910e+17	11.785714	10.0	19104.410714	6891.625000	1.267857	0.516533	11.428571
doggo, pupper	7.688683e+17	11.800000	10.0	15664.200000	5482.400000	1.200000	0.819709	7.600000
floofer	7.665514e+17	12.000000	10.0	13206.000000	4968.714286	1.142857	0.609204	7.857143
pupper	7.220454e+17	10.976471	10.0	7551.288235	2420.911765	1.288235	0.583529	9.705882
puppo	8.124018e+17	12.000000	10.0	22657.181818	6958.500000	1.454545	0.637540	12.909091

**Figure 3: Ratings by Dog Stage**