

Introduction

WeRateDogs (@dog_rates) is a Twitter handle that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerator is generally always greater than 10.

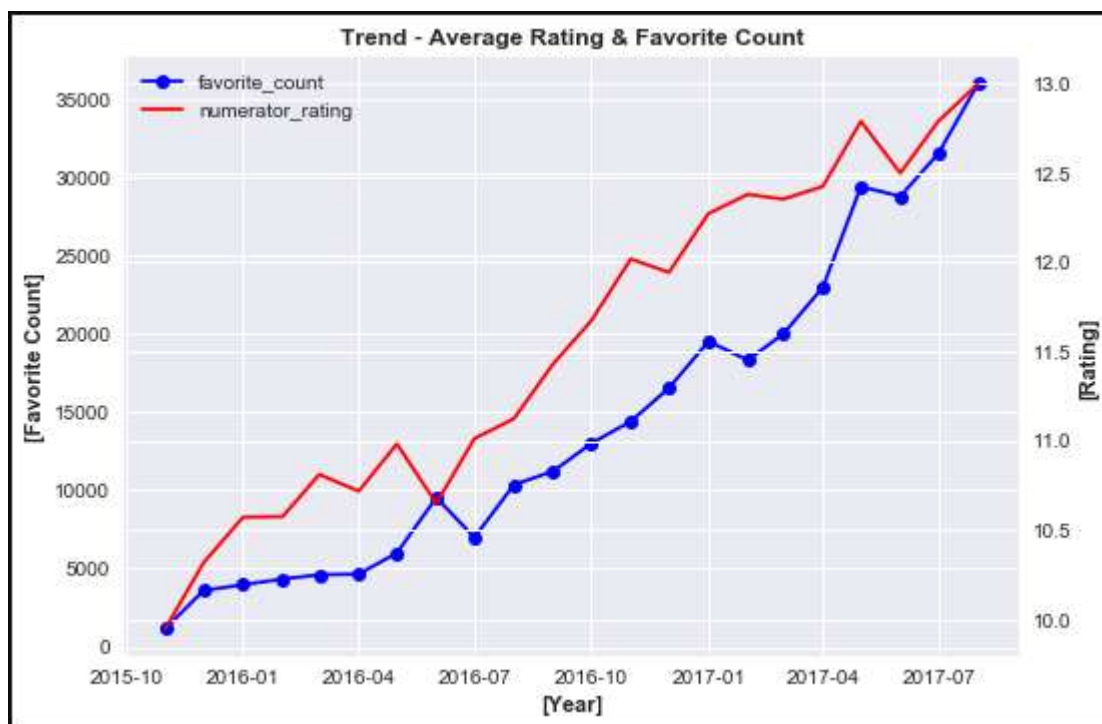
The Insights reported in this document have been arrived after data wrangling and analysing the following artefacts:

1. The enhanced twitter archives of WeRateDogs - `twitter_archive_enhanced.csv`
2. Additional relevant data downloaded from Twitter using Twitter API.
3. The tweet image predictions file (`image_predictions.tsv`) which contains a prediction of the breed of dog (or other object, animal, etc.) present in each tweet picture.

Insights

Insight # 1

- The average 'numerator ratings' per month for the dogs have gradually increased month over month - Trending up This could mean that ratings may have become a bit liberal as the time progressed.
- Also there is a change in strategy to start numerator rating from 10 onwards after the first month.
- The favorite_count per month also look to be generally increasing month over month



Act Report - Project: WeRateDogs

Insight # 2

- The top 5 rated dog breeds are (starting highest) - Pembroke ,Samoyed, chow, Golden Retriever, Cardigan
- The bottom 5 rated dog breeds are (starting lowest)- Maltese dog, Pug ,Shetland sheepdog, German Shepherd, Chihuahua

Top 5 dog breeds on the basis of their average numerator ratings/dog breed

```
In [259]: dfdog_breed_num.head(5)
```

Out[259]:

	dog_breed	favorite_count	numerator_rating	tweet_id
42	Pembroke	13,212.84	11.85	68
48	Samoyed	13,147.50	11.73	30
69	chow	11,558.65	11.73	26
77	golden_retriever	12,131.38	11.67	114
10	Cardigan	15,748.17	11.67	12

Bottom 5 dog breeds on the basis of their average numerator ratings/dog breed

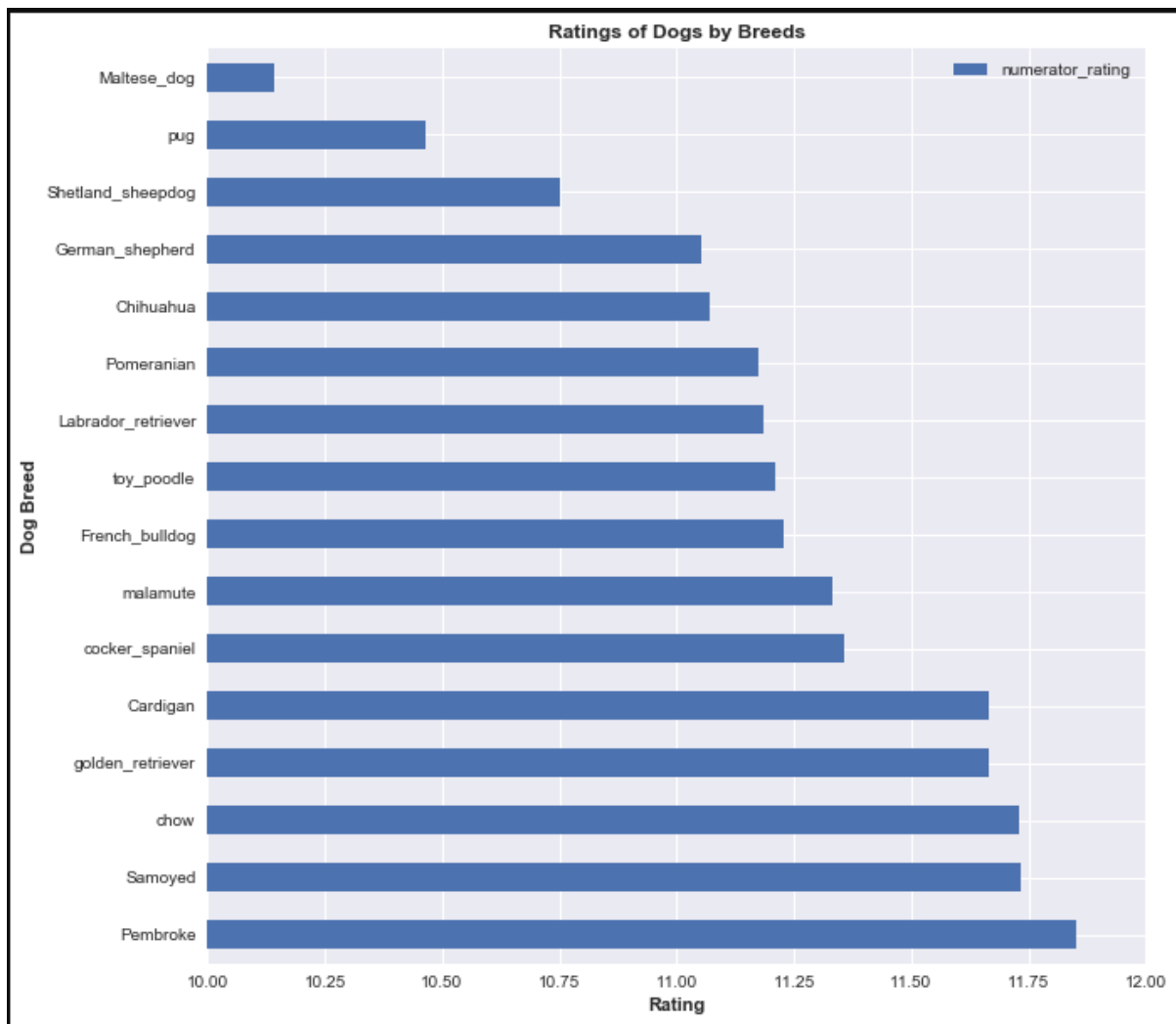
```
In [260]: dfdog_breed_num.tail(5)
```

Out[260]:

	dog_breed	favorite_count	numerator_rating	tweet_id
12	Chihuahua	10,219.02	11.07	42
20	German_shepherd	10,096.37	11.05	19
49	Shetland_sheepdog	8,597.67	10.75	12
88	pug	5,268.10	10.46	41
35	Maltese_dog	2,970.43	10.14	14

Act Report - Project: WeRateDogs

Visual Plot



Insight 3

- The top 5 most liked dog breeds are (starting highest) - French Bulldog , Cardigan, Labrador Retriever, Pembroke and Samoyed
- The bottom 5 liked dog breeds are (starting lowest)- Maltese dog, Pug ,Toy Poodle, Pomeranian and Shetland Sheepdog
- Pembroke, Samoyed and Cardigan are common in the top 5 ratings and Favorites
- Maltese dog, Pug, Shetland Sheepdog are common in the bottom 5 ratings and Favorites

Top 5 breeds on the basis of Favorite counts in Twitter / Dog breed

```
In [264]: dfdog_breed_fav.head(5)
```

Out[264]:

	dog_breed	favorite_count	numerator_rating	tweet_id
19	French_bulldog	18,595.64	11.23	22
10	Cardigan	15,748.17	11.67	12
31	Labrador_retriever	13,625.23	11.18	65
42	Pembroke	13,212.84	11.85	68
48	Samoyed	13,147.50	11.73	30

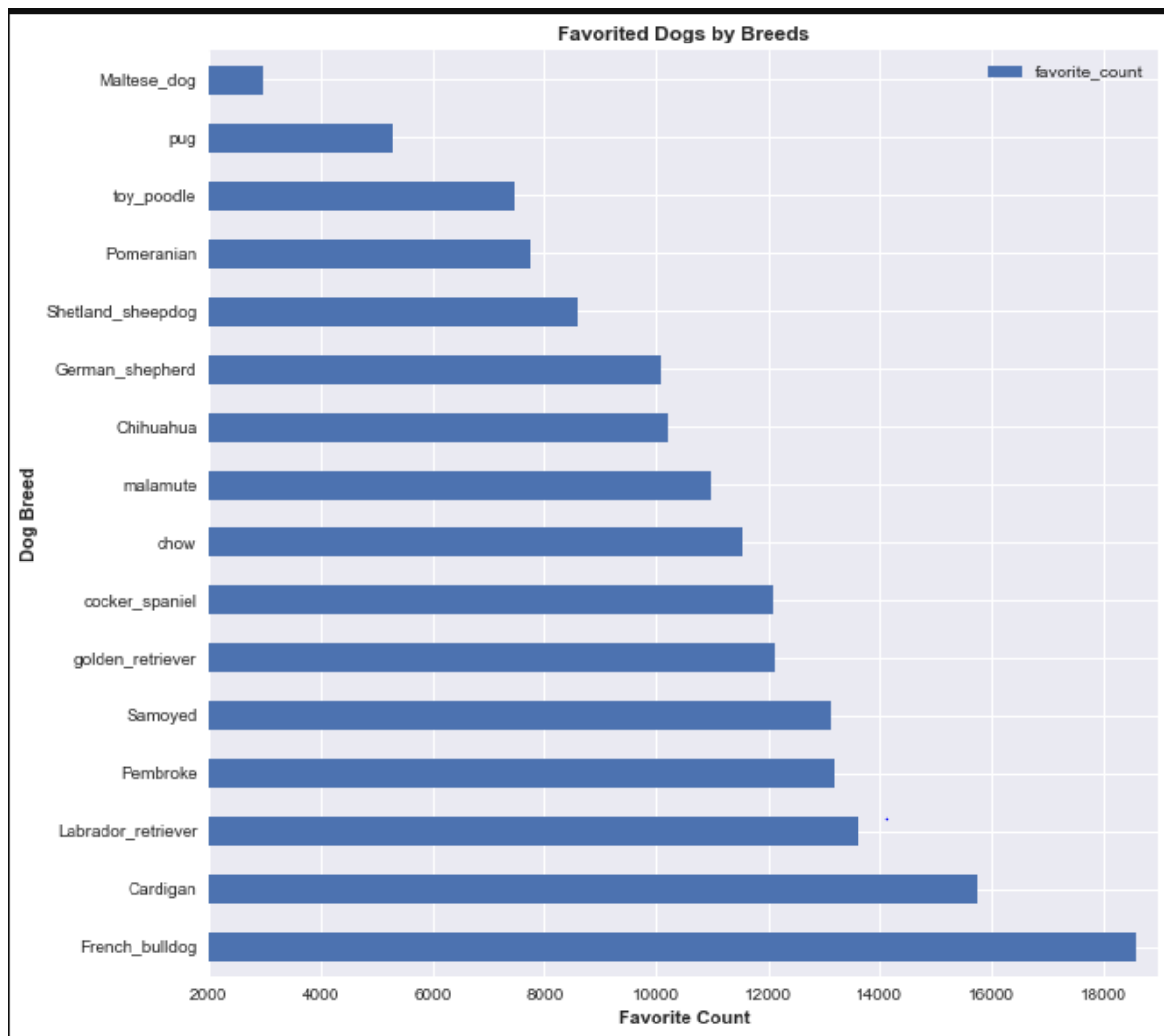
Bottom 5 breeds on the basis of Favorite counts in Twitter/ Dog breed

```
In [265]: dfdog_breed_fav.tail(5)
```

Out[265]:

	dog_breed	favorite_count	numerator_rating	tweet_id
49	Shetland_sheepdog	8,597.67	10.75	12
43	Pomeranian	7,746.41	11.17	29
93	toy_poodle	7,476.42	11.21	24
88	pug	5,268.10	10.46	41
35	Maltese_dog	2,970.43	10.14	14

Visual Plot



VISUALIZATION

Correlation Heat map

- This figure shows the correlation between different features of the Project data set
- We can see a strong correlation (0.92) between retweet count and favorite count
- There doesn't seem to be a strong correlation between the numerator rating and favorite count and retweet count

1. Correlation Heat map Plot



2. Validating the correlation between Average Rating and Favorite Count

