

act_report

June 26, 2022

1 Analysis Report

1.0.1 written by: Toluwalase Tawak

This report summarises the insights and displays the visualization(s) produced from my wrangling processes carried out on the Twitter Archive data from the @RateDogs account.

1.1 ## Introduction

For this project, I worked with three dataset provided by Udacity. They each contained different information needed to carry out analysis and reporting.

The first dataset was a csv file named `twitter_archive_enhanced`, it contained information about 2356 tweets and was downloaded manually.

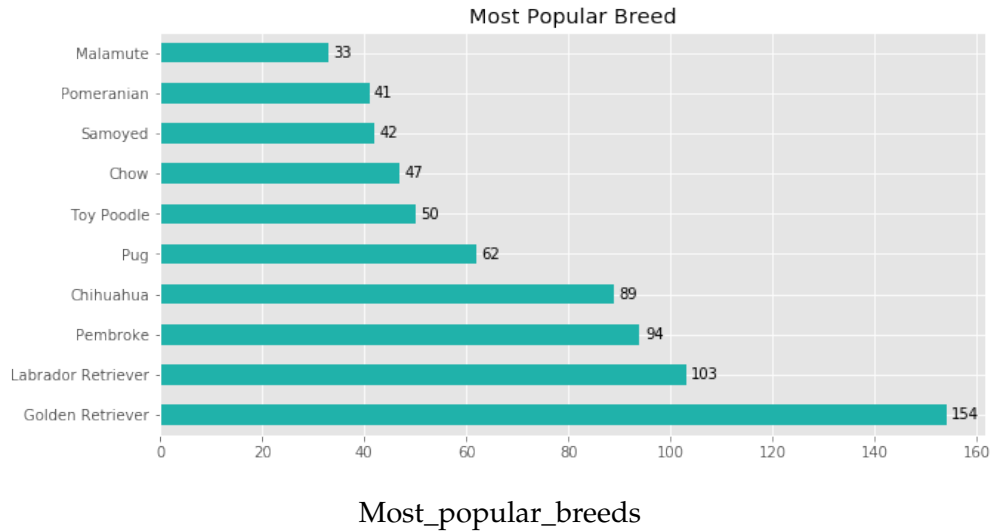
The second dataset was a tsv file named `image_prediction`, a url to the data was provided by udacity server which was I used to programmatically downloaded the file. It contained 2075 predictions classifying dogs by their breeds using pictures provided with tweets.

The third dataset, I downloaded manually as txt file containing JSON of format of tweet informations. It contained extra information about tweets like the retweet count, favorite count recieved for 2357 tweets.

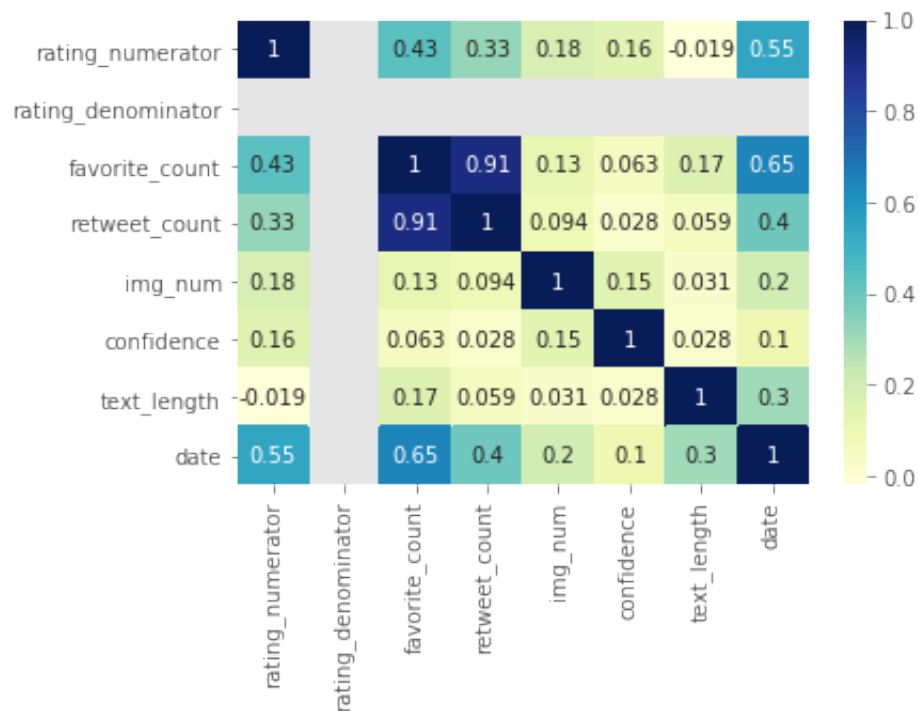
The datasets were then assessed visually and programmatically for quality and tidiness issues. These issues were addressed and corrected, a merge was carried on the three datframes to create one master dataset which was then used to carry out investigative analysis.



WeRateDogs



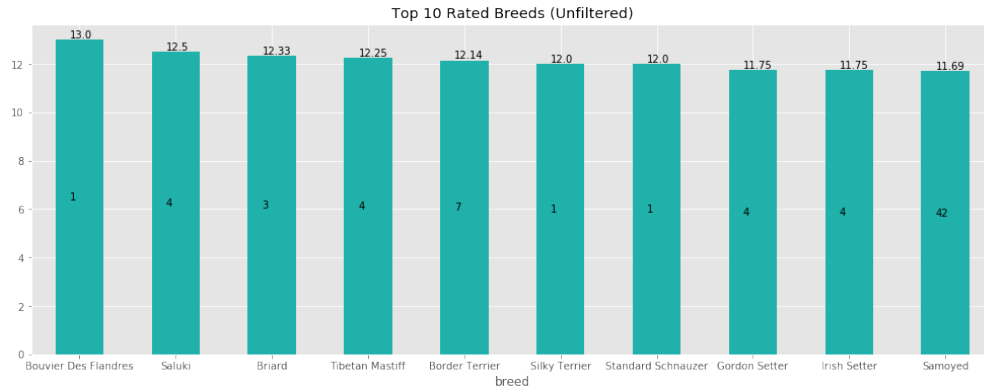
1.2 ## Analysis and Visualisations



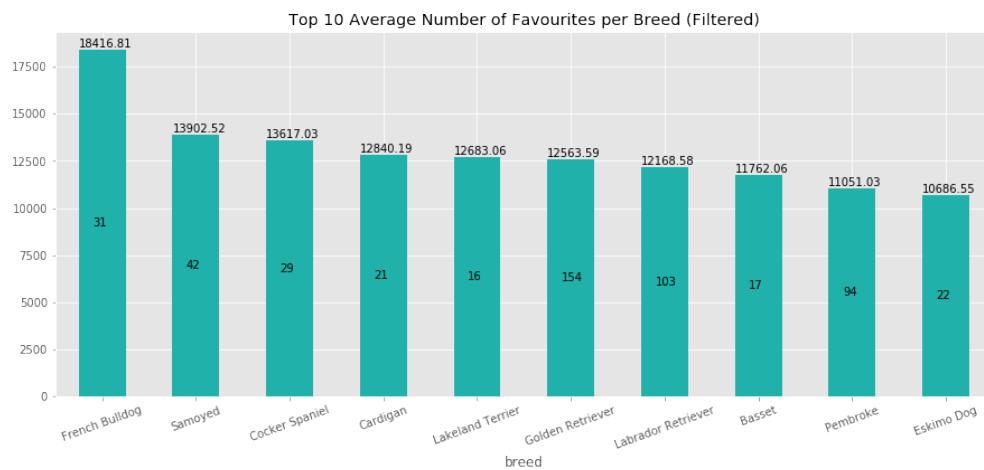
* We can see from the chart that there seems to be correlation between some attributes of our data. The strongest correlation can be observed between favourite count and retweet counts.

* We can also see that the date column seems to have a relationship with the favourite count, rating numerator and to a much lesser extent retweet count and length of tweets.

- The visual above shows us the popularity of breeds or occurrence of breeds rated by the twitter account.
- Terriers were the most talked about dogs
- Investigation shows that 6 of these breeds are small sized dogs



Top_Rated_Unfiltered



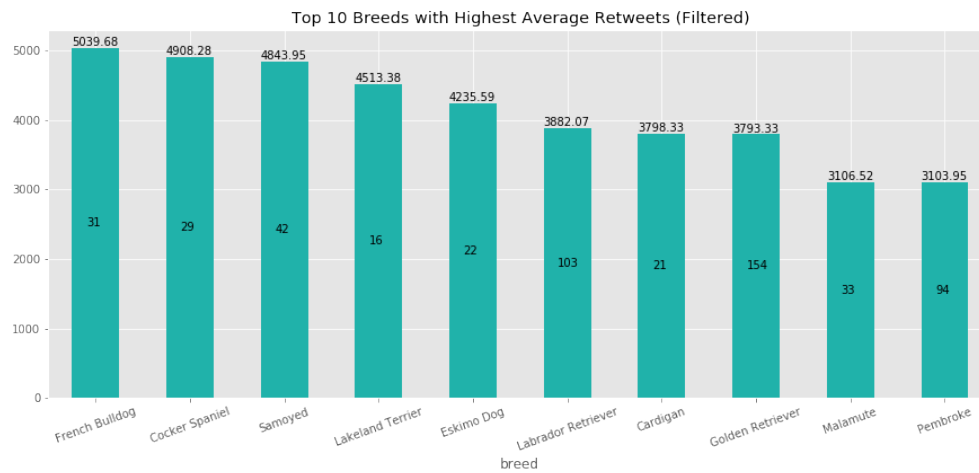
Average_Favourites_Filtered

- The highest rated breed for popular(breeds that occur more than 14 times) breed is the Samoyed.
- Golden Retriever which happens to be the most commonly tweeted dog is the second highest rated dog.
- 5 of our most common breeds within the time frame of our entire data also happen to be among the Top 10 rated breeds.
- The French Bulldog and the Cocker Spaniel, small sized dogs, are two of the top 3 popular breeds for twitter users engagemets.

1.3 ## Insights

These includes other insights not shown in this summarised report. These can be found in the Jupyter notebook used to carry pou the wrangling and analysis process.

1. The handler of the this twitter account like to use the abbreviation "af" a lot. WHICH means he/she liked to emphasize their description or sentiments for any dog they were rating.



Avg_Retweets_Filtered



WordCloud

2. Of the four different classification of dogs, Pupper was the handler's favourite description for dogs. From the 'dogtionalary', these dogs are small or young dogs.
3. In the time period under consideration, the maximum number of characters for a tweet was 140. Therefore, the handler used a lot of long tweets, between 100 and 140 characters for most of their tweets.
4. The handler was very generous with ratings, giving most dogs a numerator 10 or more.
5. The two most popular breeds rated, were both Retrievers. If people sent their dog's pictures to be rated, most of the dogs owners who are aware of the account owned Retrievers (Golden and Labrador). If the handler found the pictures on their own, is it easier to find pictures of retrievers than other breeds?
6. It appears the handler has a bias towards the Golden Retriever, as it is also closely the dog with second highest rating for popular dogs. The Labrador Retriever is also in the top 10 for popular dogs with the highest ratings.
7. The most popular dogs also had the most total engagements in terms of Favourites and retweets.
8. Followers of the accounts seemed to have lots of love for the French Bulldog, Cocker Spaniel and the Samoyed as they were the most popular dogs that received the most likes or retweets on average.
9. The Bedlington Terrier and Saluki breeds are probably the breed followers of the account liked the most. They received the most total engagements on average even with their lack of popularity. (All though these engagement values could be skewed due to just one of their pictures receiving very very large number of engagements.)

```
In [5]: !jupyter nbconvert --to pdf wrangle_report.ipynb
```

```
[NbConvertApp] Converting notebook wrangle_report.ipynb to pdf
[NbConvertApp] Writing 22715 bytes to ./notebook.tex
[NbConvertApp] Building PDF
[NbConvertApp] Running xelatex 3 times: ['xelatex', './notebook.tex']
[NbConvertApp] Running bibtex 1 time: ['bibtex', './notebook']
[NbConvertApp] WARNING | bibtex had problems, most likely because there were no citations
[NbConvertApp] PDF successfully created
[NbConvertApp] Writing 42724 bytes to wrangle_report.pdf
```

```
In [ ]:
```