

ACT REPORT ON DATA WRANGLING PROJECT (WERATE DOG)

ANALYSIS & VISUALIZATIONS:

Introduction:

This is a report on the analysis, insight and visualization carried out after the datasets had been cleaned and merged to get an accurate representation of the information on the three datasets which include: Twitter_Archive.csv, image_predictions.tsv and the tweet.json datasets.

i.FIRST INSIGHT:

Analysis was carried out on the p1 column for dog breed to determine the most common dog breed in the dataset. And in conclusion I was able to get the 5 most common dog breeds in the dataset.

The five most common dog breeds in the dataset are as follows:

- | | |
|-----------------------|-----|
| a. Golden retriever | 150 |
| b. Labrador retriever | 100 |
| c. Pembroke | 88 |
| d. Chihuahua | 83 |
| e. pug | 57 |

ii.SECOND INSIGHT:

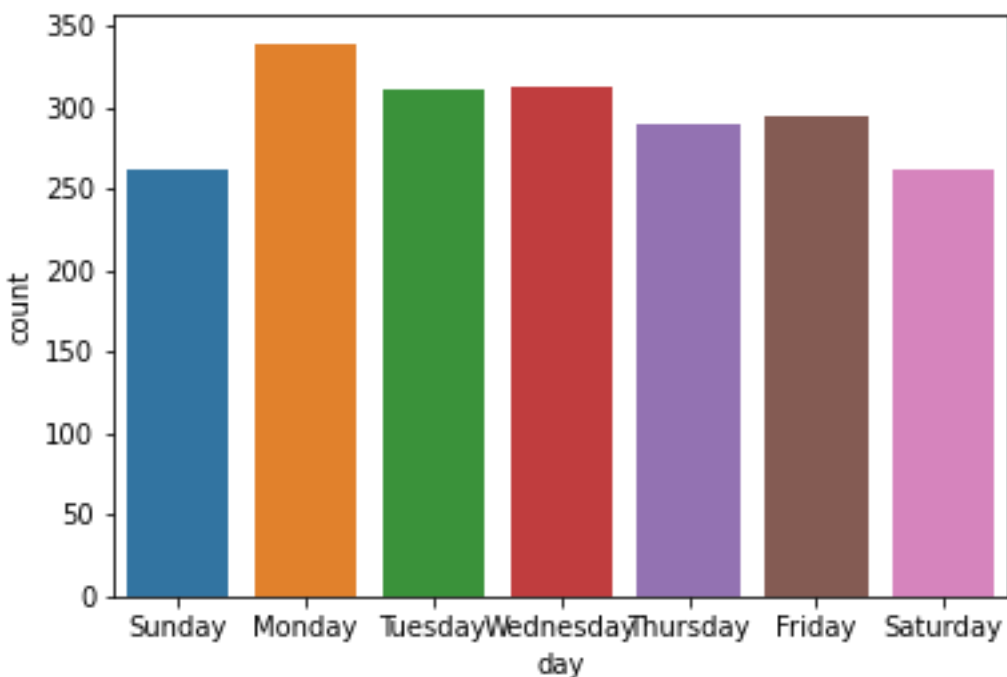
Further analysis was also carried out on the new dog_stage column to determine the most common dog stage in the dataset. Though a significant amount of dog_stage was not given according to the data I could work with. These are the most common dog stage tweeted about in the dataset:

1.NONE	1752
2.pupper	209
3.doggo	67
4.puppo	23
5.doggo,pupper	11
6.floofer	7

VISUALIZATION

i.FIRST OBSERVATION AND VISUALIZATION

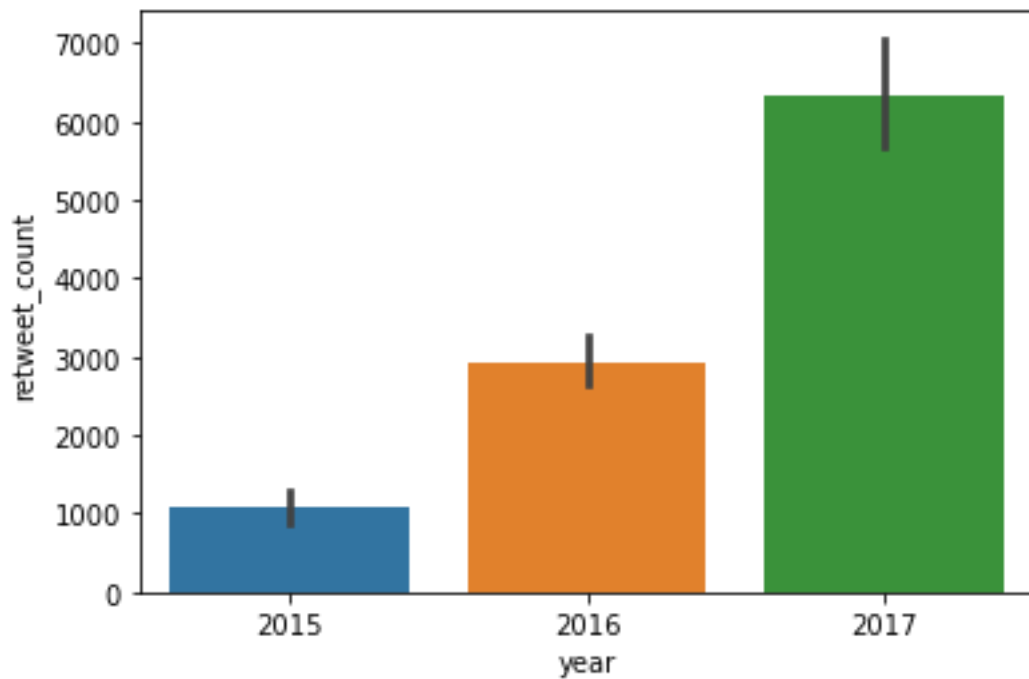
According to the days the tweets were sent. When we analyzed the tweets Mondays was the busiest day in which we had the most frequent tweets. Below are the number of tweets per day according to the dataset and the Visual representation of per day of the week.



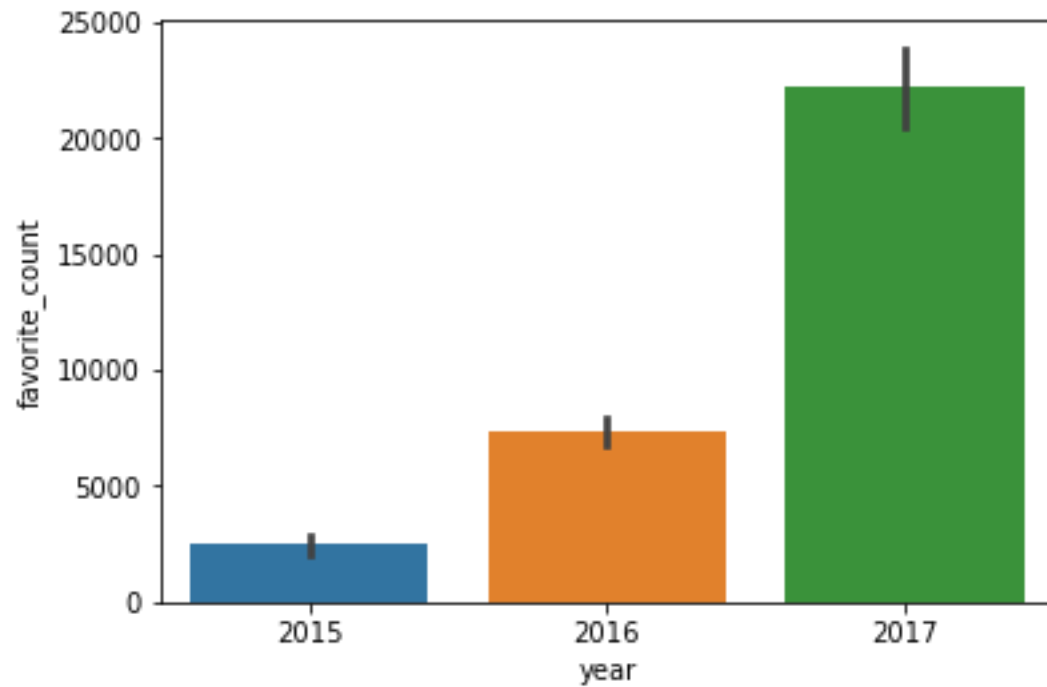
1.Monday	339
2.Wednesday	313
3.Tuesday	310
4.Friday	295
5.Thursday	290
6.Sunday	262
7.Saturday	262

ii.SECOND OBSERVATION AND VISUALIZATION

There has been a significant increase in the number of users for weratedogs based on data collected from both users that like the tweets and those that retweeted over the years from a period of 2015-2017.



Tweets based on yearly basis



Tweets based on likes(favorites)

iii.FINAL OBSERVATION AND VISUALIZATION

There is a correlation between retweets and favorites. An increase in the number of retweets led to an increase in the number of favorite tweets too.

