

# Act Report

Img1 :Front page of the twitter page of @rate\_dogs

**Introduction-** After all the wrangling and cleaning, now is the time to analyse the cleaned data and withdraw a few conclusions. We will be drawing analytical and visual conclusions of the data we collected about the twitter user id @rate\_dogs.

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, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "they're good dogs Brent."

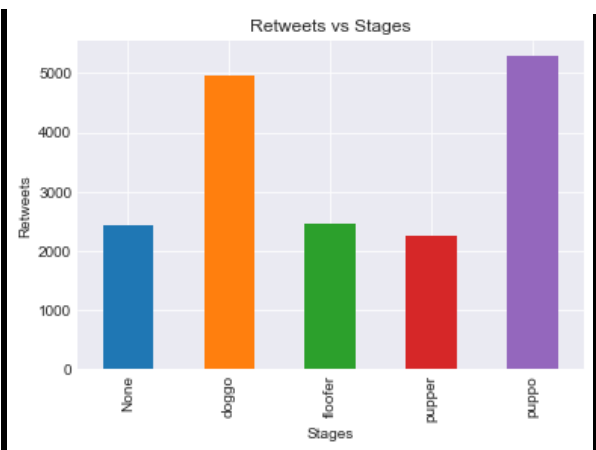
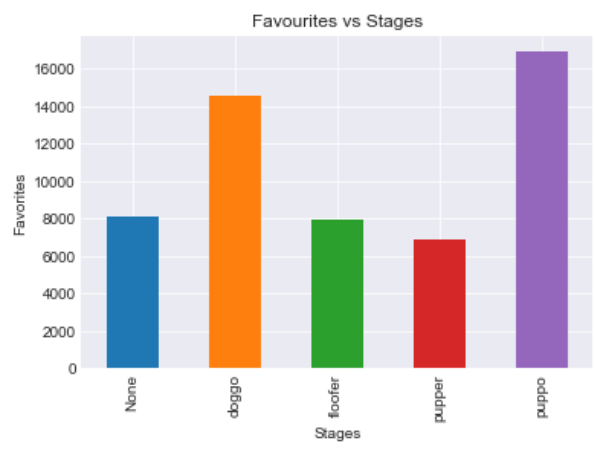
## All the Analysis Made-

I made 4 analysis from final table after the Wrangling process :

### 1. Relation between Stage of the dog in the photo and the popularity of the tweet-

The first question that came across my mind while analysing was does the stage stages(doggo,floofer,puppo,fluffer) of the dog in the tweet affect how appealing the tweet is ?

I turned out to be right as tweets with photos of puppos garnered more retweets and favourites and also gained higher rating score by the tweeter. Puppos had highers mean retweet and favourite score than all other stages.

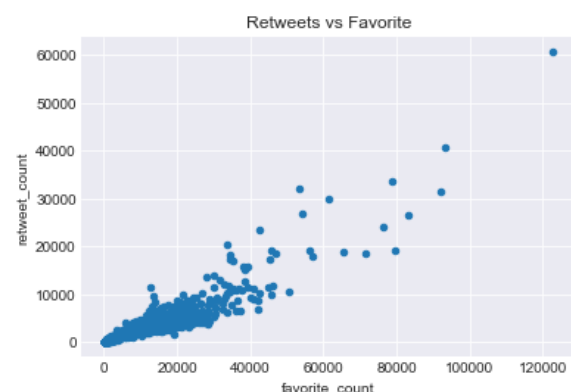
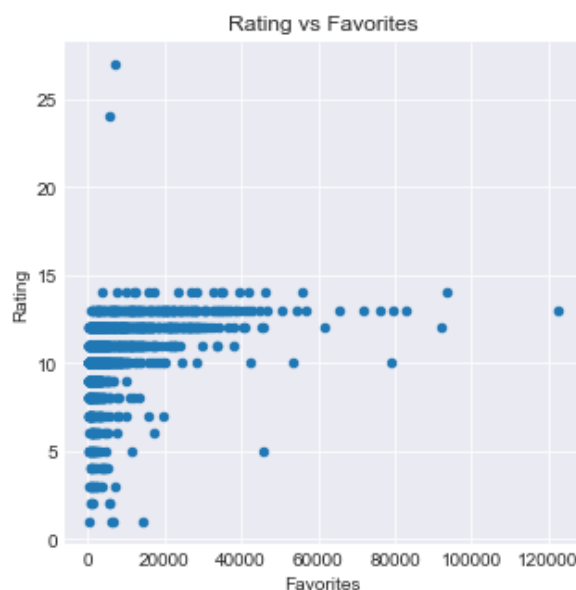




Img2: Tweet and the images of the puppo with the highest retweets and likes

## 2. Getting a relation between Ratings given in the tweet and the likes that the tweet gets:

I felt like deriving a relationship between the Ratings and Favourites to find out whether the ratings provided by the tweet manipulate the audience to like the tweet. By visual interpretation there was no specific trend which could be derived as tweets with median ratings got higher likes than tweets with higher ratings. Although images with higher rating tend to have higher number of likes



Secondly I also drew a scatter plot for retweets vs Favorites which showed positive correlation between the both thus leading to the conclusion that higher retweets result in higher likes and vice versa.

### 3. Do Names affect the how well the tweet is received?

The data in the table had information about the name of the dog derived from the text part of the tweet. But some tweets lacked the name and were simply text with funny captions. Which made me wonder if the name affects the performance of the tweet?

Surprisingly tweets with no name actually performed better in terms of retweets as tweets with no names got an average retweet count of 2573 and tweets with names got a retweet mean of 2493.

However tweets with name in them have the highest ratings and retweets but I think that was actually down to the funniness of the images.



Img 4 and 5 : Tweet with the highest ratings(left) and tweet with the highest retweets

### 4. Analysing the column gathered from Image Prediction Table:

For the Image prediction table the images were run through a neural network to get the predicted dog breed and the confident prediction.

So I analysed the confidence of each prediction and unsurprisingly had the highest mean confident prediction value.

```
In [362]: Image_predictions_analysis.describe()
Out[362]:
```

|       | p1_conf     | p2_conf      | p3_conf      |
|-------|-------------|--------------|--------------|
| count | 1299.000000 | 1.299000e+03 | 1.299000e+03 |
| mean  | 0.587034    | 1.370495e-01 | 6.144723e-02 |
| std   | 0.273638    | 1.018687e-01 | 5.202736e-02 |
| min   | 0.044333    | 1.011300e-08 | 1.740170e-10 |
| 25%   | 0.354703    | 5.438335e-02 | 1.649005e-02 |
| 50%   | 0.578120    | 1.203580e-01 | 4.953060e-02 |
| 75%   | 0.837040    | 1.982365e-01 | 9.470910e-02 |
| max   | 1.000000    | 4.676780e-01 | 2.710420e-01 |

