REPORT: ACT\_REPORT

The WeRateDogs tweets dataset contained names, ratings of dogs etc. as well as favorite counts and retweet counts of each of the tweets. As with any data, we want to gain insights by asking questions based on the data. For example, we may pose questions like “What are the popular dog names?” or “Which tweet had the highest rating?” etc. Analysis of the data produced the following insights:

* The average dog rating numerator was 13.3.
* The tweet with ID 744234799360020481 involving a doggo had the greatest number of retweets (79,515). The dog also had a rating of 13. Also, this tweet (822872901745569793) involving a *puppo* had the most favorite count (132,810)
* The most popular dog names found in the dataset were Lucy and Charlie (figure 1).
* The higher the rating, the higher the favorite counts.

**Table 1: Descriptive statistics of dog ratings, favorite counts and retweet counts**

|  |  |  |  |
| --- | --- | --- | --- |
| Parameters | Dog Rating (numerator)a | Retweet count | Favorite count |
| **count** | 2094.00000 | 2094.000000 | 2094.000000 |
| **mean** | 13.32999 | 2687.651385 | 8463.776982 |
| **std** | 48.63494 | 4615.512215 | 11797.592466 |
| **min** | 0.00000 | 0.000000 | 52.000000 |
| **25%** | 10.00000 | 591.250000 | 1834.250000 |
| **50%** | 11.00000 | 1311.000000 | 3897.500000 |
| **75%** | 12.00000 | 3121.250000 | 10810.000000 |
| **max** | 1776.00000 | 79515.000000 | 132810.000000 |

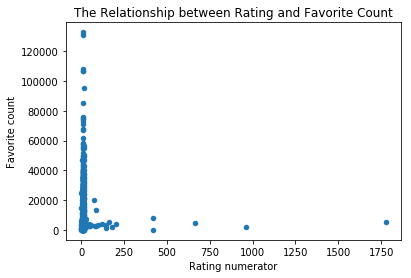
1. The denominator is omitted since it’s a constant value (10).



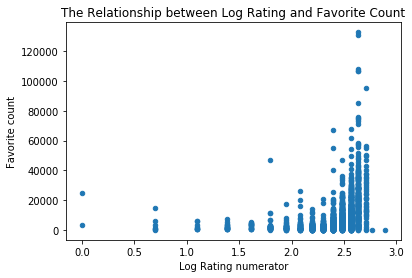
**Figure 1: The Most common dog names among WeRateDogs’ Tweets.**

**The relationship between dog rating and favorite counts**

We investigate the relationship between the dog rating and tweets favorite counts (figure 2) as well as between the log rating and the favorite count (figure 3). We see that the plot below is affected by outliers. These outliers were usually made up of the ratings of a group(s) of dogs. The relationship between the favorite count and log rating provided a clearer picture (figure 3). It appeared that favorite count increased with the rating although the trend seemed like a curved one



**Figure 2: The relationship between the dog rating and tweets favorite counts**



**Figure 2: The relationship between the log rating and tweets favorite counts**