

Analysis

Some conclusions I reached upon analysis the WeRateDogs tweet Data after combining all the three files:

a) Correlation between retweets, ratings_num and favorites

| | favorite_count | retweet_count |
|----------------|----------------|---------------|
| favorite_count | 1.00000 | 0.92633 |
| retweet_count | 0.92633 | 1.00000 |

With this correlation we can say that there is strong correlation between favorites and retweets.

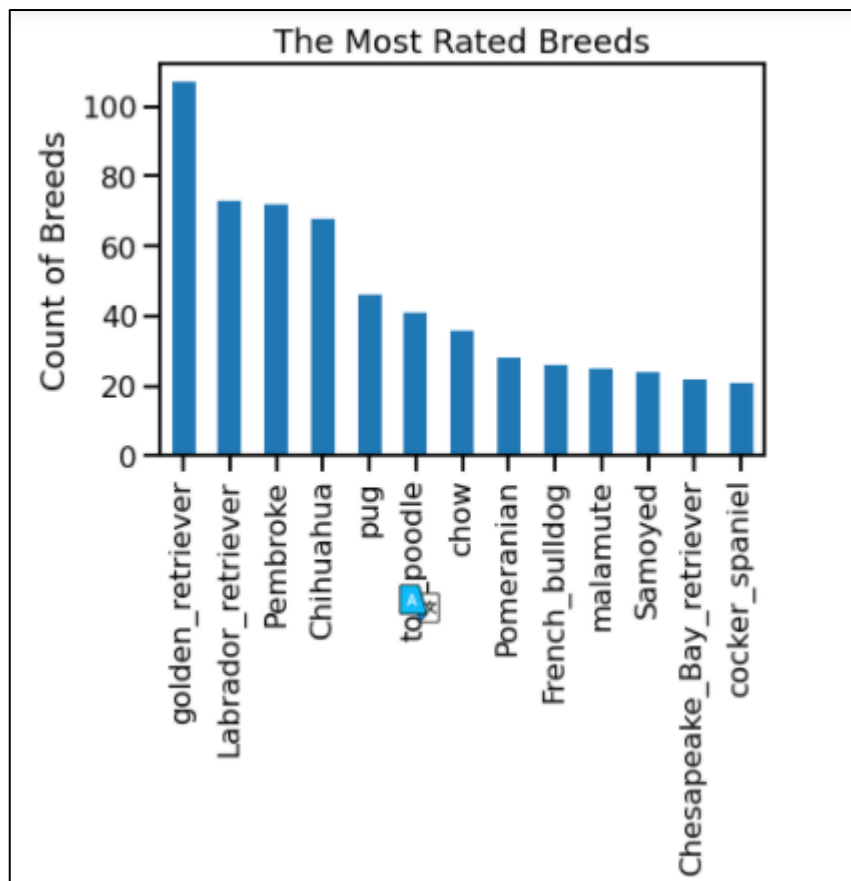
b) From analyzing stage of dogs, along with the ratings we concluded that puppies are the one who got the lowest ratings among all other breed of dogs whereas floofer are the one who got minimum numerator ratings is 10 so we can say that they are consistently good rated dogs.

| | count | mean | std | min | 25% | 50% | 75% | max |
|---------|--------|-----------|-----------|------|------|------|------|---------|
| stage | | | | | | | | |
| None | 1190.0 | 12.170294 | 51.223705 | 2.0 | 10.0 | 11.0 | 12.0 | 1776.00 |
| doggo | 36.0 | 12.083333 | 1.480830 | 8.0 | 11.0 | 12.0 | 13.0 | 14.00 |
| floofer | 5.0 | 11.600000 | 1.140175 | 10.0 | 11.0 | 12.0 | 12.0 | 13.00 |
| pupper | 122.0 | 10.998033 | 7.630242 | 3.0 | 10.0 | 10.5 | 11.0 | 92.49 |
| puppo | 15.0 | 11.800000 | 1.264911 | 9.0 | 11.5 | 12.0 | 13.0 | 13.00 |

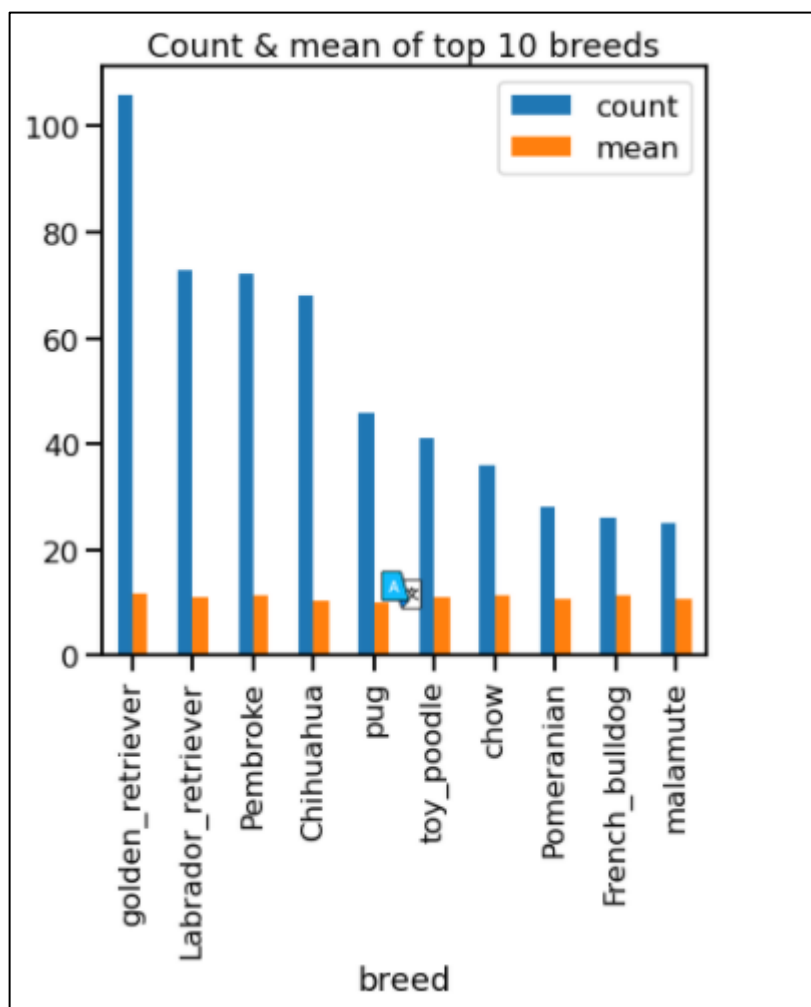
c) Combining dog stages with timestamp we can say that floofer is not a new term, they were first introduced on Jan 2016. Combining this with minimum numerator ratings which is 10 we can say that floofers got consistently good ratings.

| | count | unique | top | freq |
|---------|-------|--------|---------------------------|------|
| stage | | | | |
| None | 1191 | 1191 | 2015-12-19 16:47:58+00:00 | 1 |
| doggo | 36 | 36 | 2016-08-23 21:09:14+00:00 | 1 |
| floofer | 5 | 5 | 2016-05-21 00:50:46+00:00 | 1 |
| pupper | 122 | 122 | 2015-12-06 00:17:55+00:00 | 1 |
| puppo | 15 | 15 | 2016-06-03 01:07:16+00:00 | 1 |

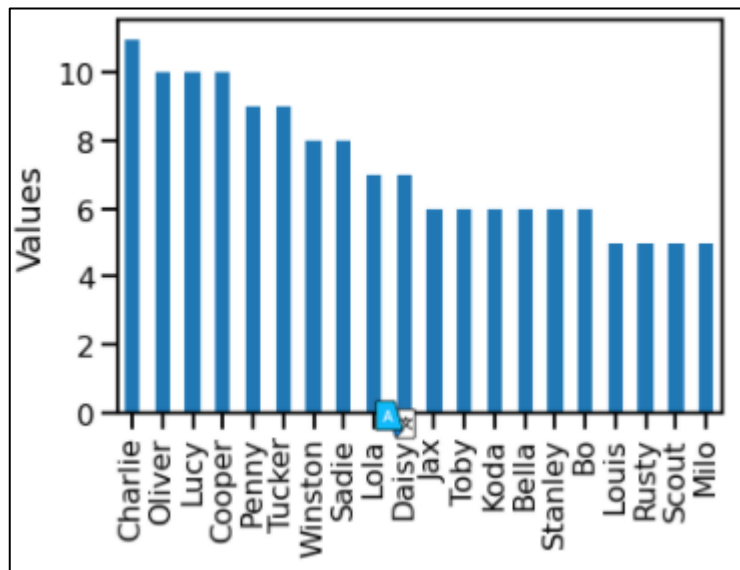
D) From analyzing top ten most rated dogs we analyzed golden retriever are the most rated dogs among all other breeds of dogs whereas cocker spaniel are the least rated dogs.



E) Combining top ten most rated dogs with numerator ratings we got to see that chow breed has the highest mean ratings whereas pug breed got the lowest ratings among top 10 most rated breeds.



F) In top twenty popular dog names, 'Charlie' is the most favorite name whereas 'Milo' is the least favorite name



G) From analyzing Favorites and retweets trend over time, we realized increment in favorites and retweets over time. We can say that there is increment in popularity of "WeRateDogs" over time.

