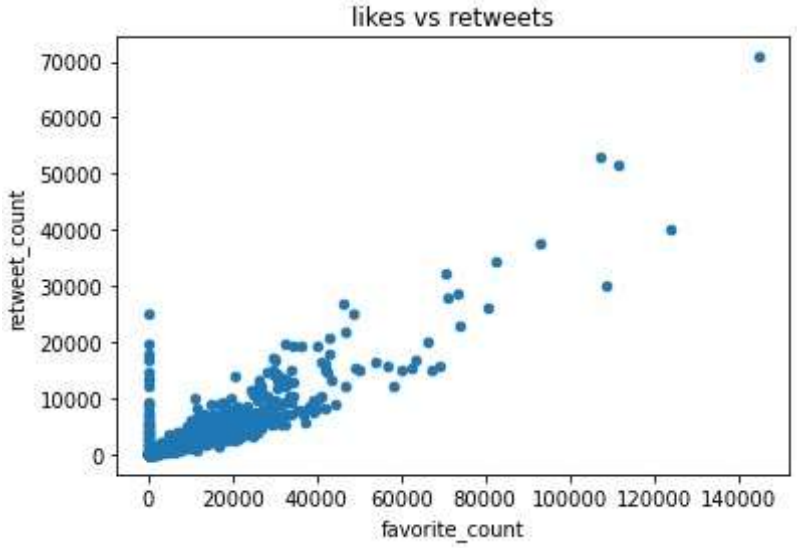
Report: act_report

- Create a **250-word-minimum written report** called "act_report.pdf" or "act_report.html" that communicates the insights and displays the visualization(s) produced from your wrangled data. This is to be framed as an external document, like a blog post or magazine article, for example.

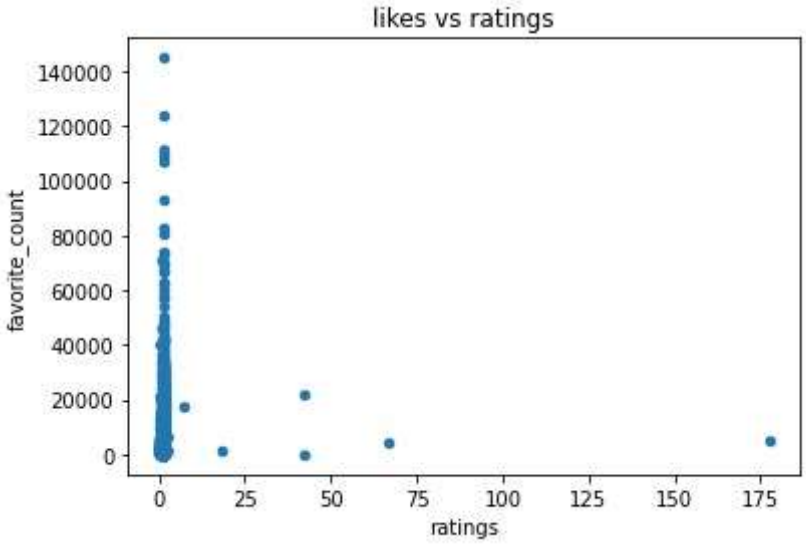
Analyzing tweets of dog ratings from the WeRateDogs Twitter Archives

We Rate Dogs is a twitter account that rates dogs with humorous comments. It has quite a lot of followers. They group dogs into four categorizes: doggo, floofer, pupper, and puppo, and uses a unique rating system. I wanted to investigate if there is any correlation between dog ratings, number of likes and retweets. In other words, I wanted to know if the number of likes a tweet gets is probably due to a dog rating. So, I gathered tweets from November 15, 2015 to August 1, 2017.

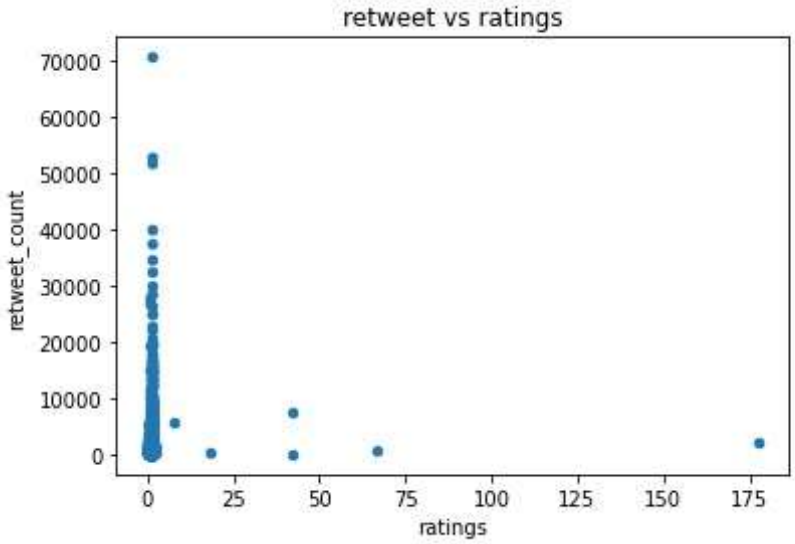
After cleaning the data, I found that there is a strong correlation (0.88 pearsonr coefficient) between the number of likes and retweets which is expected. We can also see this in the scatter plot below



This means people tend to retweet tweets that they like. However, there is almost no correlation (0.014 pearsonr coefficient) between dog ratings and the number of likes which is quite evident from the scatter plot below



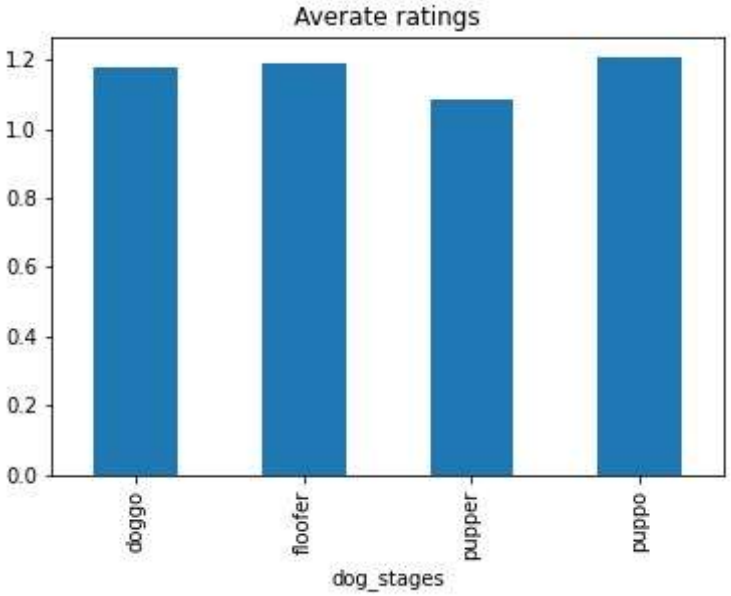
There is also almost no correlation between dog ratings and the number of retweets.



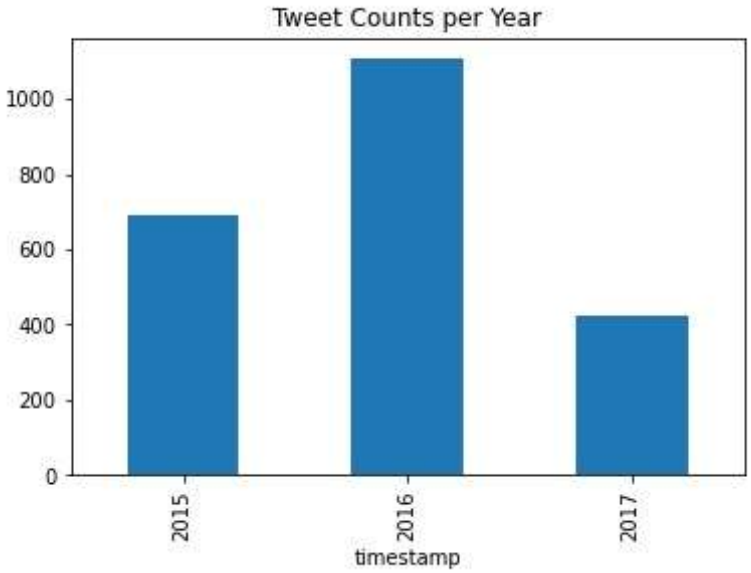
This means that there might be other reasons why people like tweets with dog ratings. Next, I wanted to investigate which of the dog stages gets the highest average rating. These are the mean of the ratings for each dog stage based on the data I gathered

```
dog_stages
doggo      1.178409
floofer    1.188889
pupper     1.082969
puppo      1.204167
Name: ratings, dtype: float64
```

which is visualized in the bar plot



It looks like puppos have the a slightly higher average rating compared to the other dog stages. Lastly, I wanted to check the total number of ratings made per each year of the data collected. My findings is shown in the bar plot below



From here, we can see that there were more tweets of dog ratings made in 2016.