ACT REPORT

WeRateDogs is a twitter account that rates people's dogs with a humorous comment about the dog. The WeRateDogs twitter archive contains basic tweet data for all 5000+ of their tweets posted between November 15, 2015 and August 1st, 2017. The retweet and favorite counts were not included in the archive and so I needed to obtain them from the twitter account.

Wrangling

Before analyzing the data and generating insights from it, I had to gather data, assess it and then clean it to ensure that there would be no issues with its quality or structure.

Insights

- **The rating with the highest favorite count is 13.5 while the rating with the rating with the lowest favorite count is 11.26.**

I used the groupby function to group the columns based on the rating_numerator column. I then plotted a bar chart to see the number of favorite_count(likes) each rating_numerator value had. The ratings on the x axis and the favorite count on the y axis.

- **The pupper is the dog_stage that was posted about the most.**

	1744
pupper	221
doggo	72
puppo	23
floofer	9
doggo, pupper	9
doggo, puppo	1
doggo, floofer	1

I used the value_counts function for the dog_stages column. This revealed that the pupper dog stage was posted about the most.

- **The dog stage with the most likes is the doggo, puppo.**

I used the groupby function to group the columns based on the dog_stages column. I then plotted a bar chart to see the number of favorite_count(likes) each dog stage value had. The dog stage on the x axis and the favorite count on the y axis.

Limitation

It is important to note that the conclusions drawn from this analysis are limited. This is because a significant portion of the original data had been removed during the cleaning process. Therefore, I worked with a much smaller sample size and as such does not necessarily represent the population.