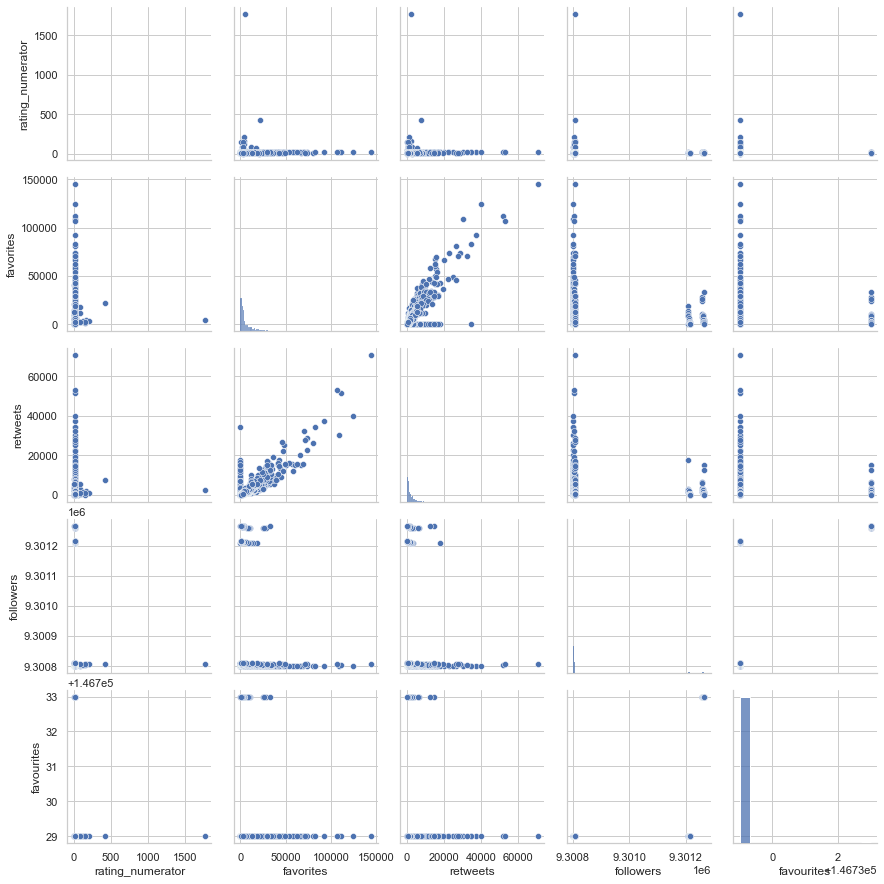
**Analysis Report**

**Ednalyn C. De Dios  
D309, WGU  
07/09/2022**

## Pairplot

I used a pairplot to compare how variables compare with each other.

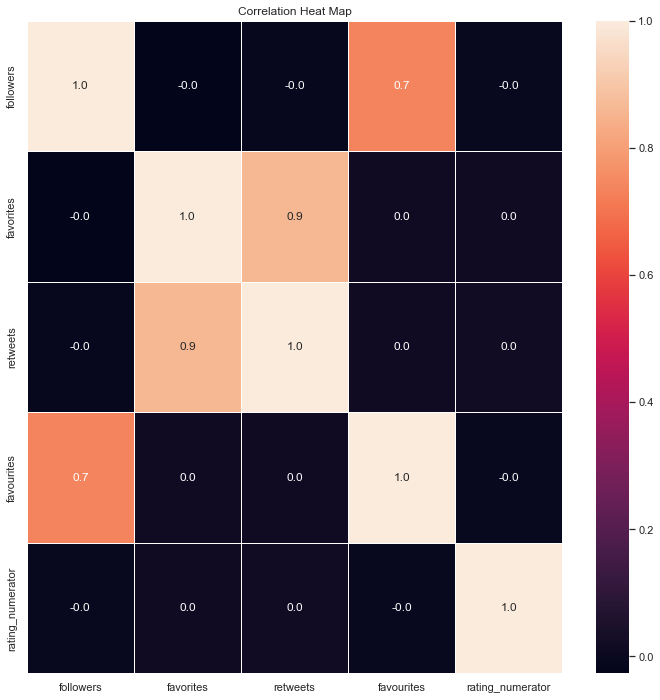
There is strong positive correlation between favorites and retweets. This is helpful to know if we are to isolate truly independent variables to avoid multicollinearity.



## Heat Map

Another visualization I used to compare the variables to each other is the heat map.

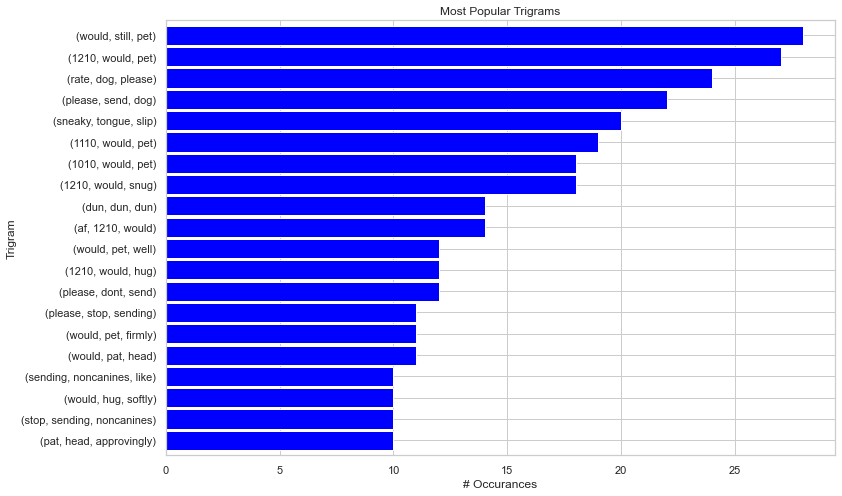
Below, we see a strong correlation between favorites and retweets, corroborating the findings above. Also, we observe that favourites and retweets are also correlated albeit not as strong as the previous one.



## NLP

A simple plot of the most popular trigrams is shown below.

From cursory examination, we can see that the textual data needs most cleaning. The graph also suggests that there are possible word combinations to investigate like "would still pet", "rate dog please", and "please don't send."



# Conclusion

The WeRateDogs is fun to play with. Since it has textual freeform data, it also presents a good opportunity to try out some NLP techniques.