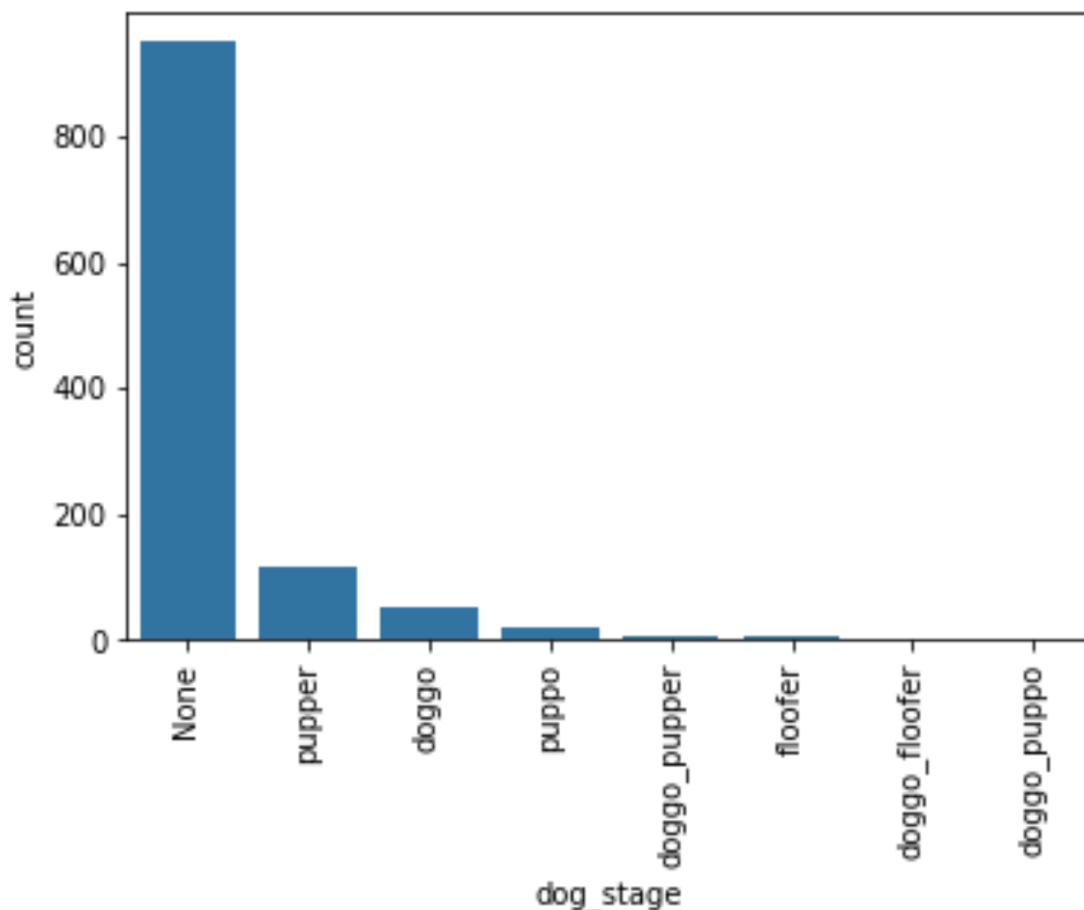


Analyzing the Wrangled and Cleaned WeRateDogs' Twitter Data

WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. In this project, this twitter data of WeRateDogs was gathered, assessed and cleaned. After cleaning, the dataset has 1153 tweets. Retweets, tweets with rating less than 1 and tweets without images have been removed from the dataset. The following analysis has been made on the cleaned dataset.

1. Finding the most used dog stage in tweets

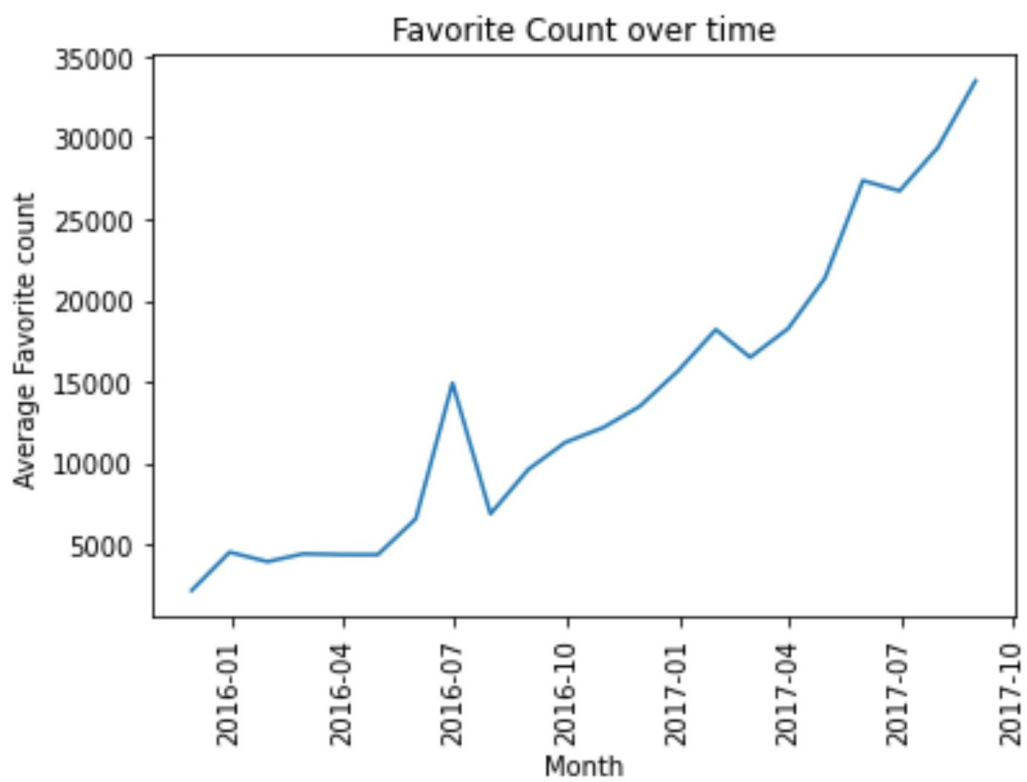
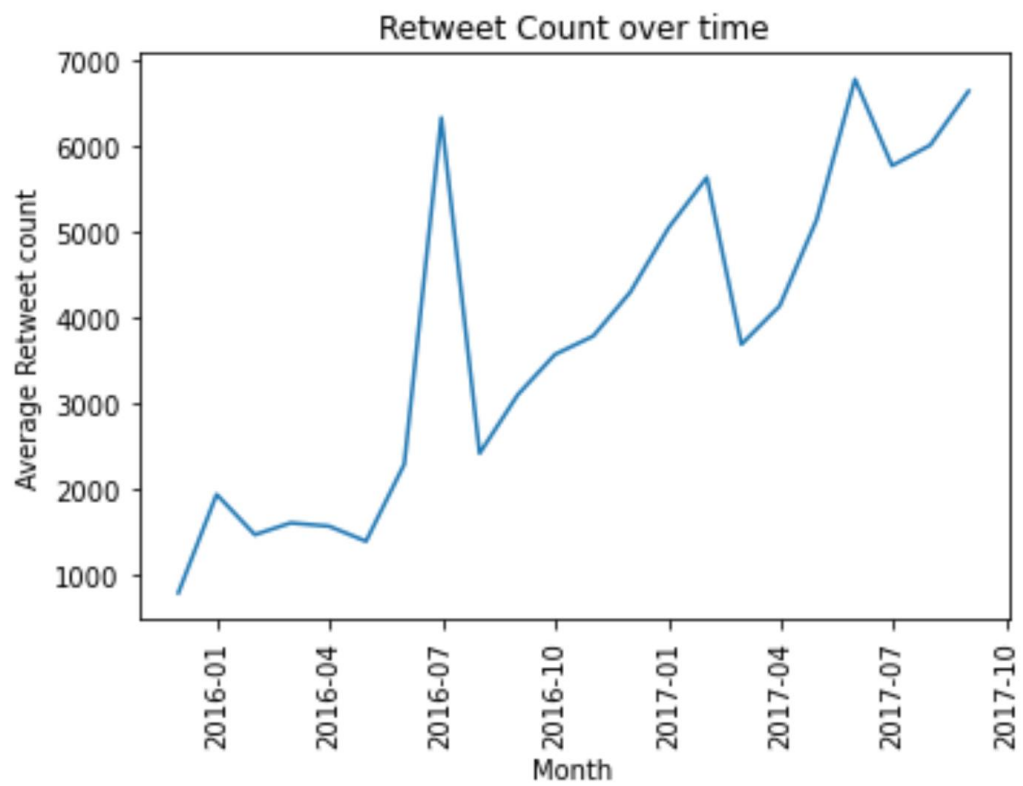
It is observed that majority of the tweets in the wrangled and cleaned dataset did not mention the dog stage. Among the tweets that did mention, pupper was the most commonly used. This was followed by doggo and puppo.



2. Trends in retweet count and favorite count

The average retweet count and favorite count for each month have been calculated and plotted. It is observed that both retweet count and favorite count peaked in July, 2016. After July, 2016, favorite decreased only for few months and then showed an increasing trend. Retweet count too showed an increasing trend but it was not smooth and had peaks and dips. Moreover while the

favorite count surpassed the peak in July'16 greatly, the retweet count has not increased much greater than July'16.



3. Most common top image prediction

A neural network classified the images of the dogs into breed. The top prediction was analysed. Golden retrievers is the most commonly predicted breed for the dog. This is followed by Pembroke and Labrador retriever.

