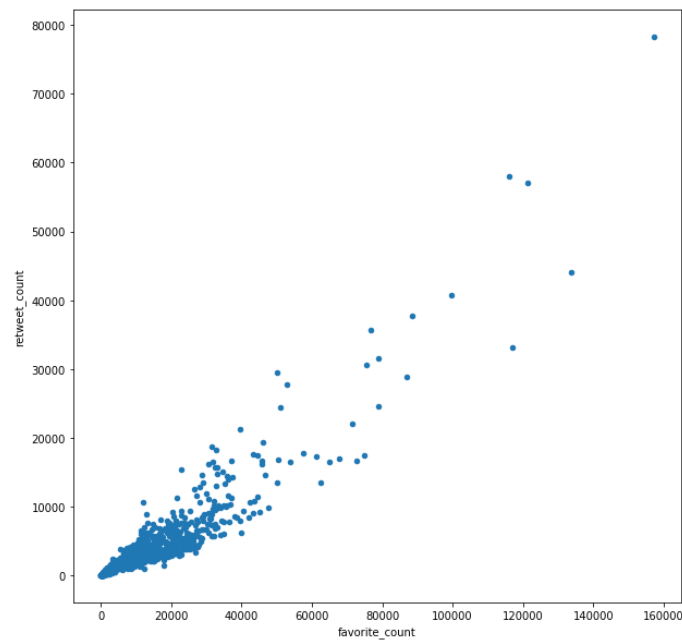
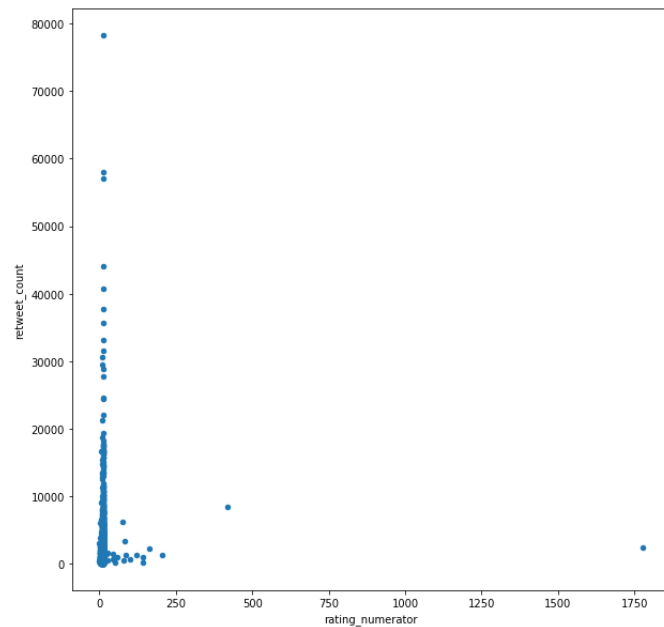


### Insights and Visualisation

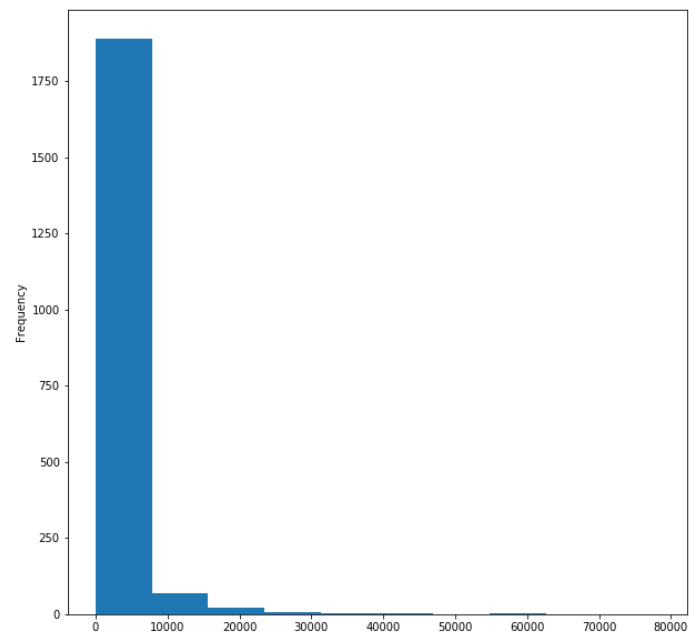
The project began with data gathering which required opening a developer account with Twitter in order to access the API. Once that was done, the next stage began easier. The assessment of the combined data was tedious. There are many issues in the data as described above. However, for this project only 8 quality issues and 2 tidiness issues were resolved. The final product has 25 columns down from 32 columns from all three tables. The data now is ready for visualisation. A total of 8 plots were made based on the findings.



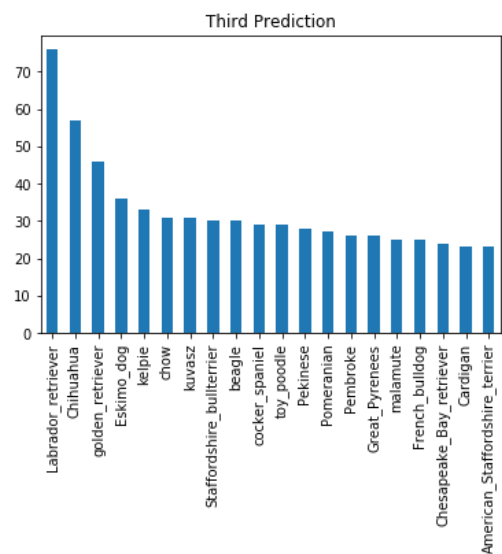
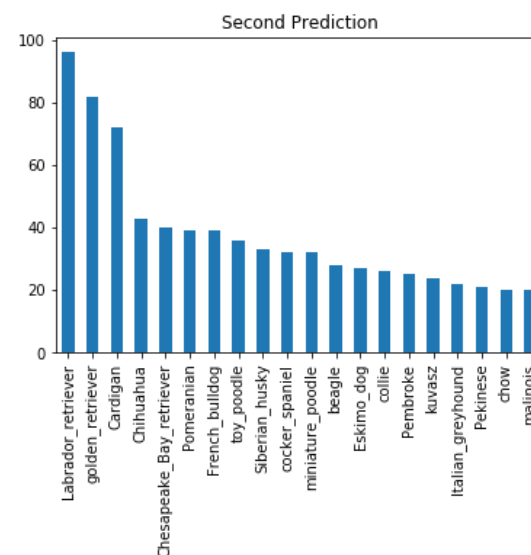
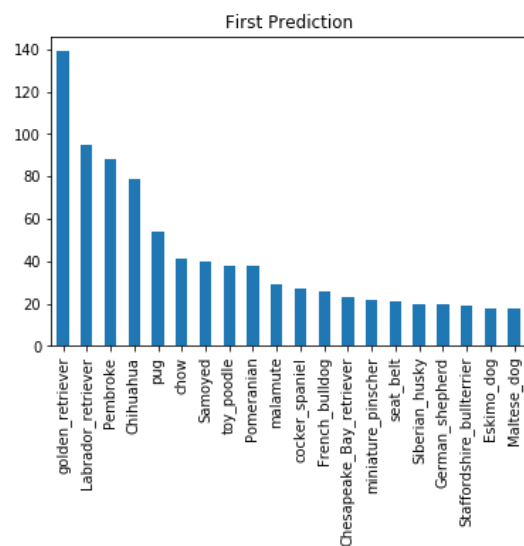
This plot shows the count of favourite against the count of retweeted tweets. The plot shows a congregation within the 40000 mark at which rate the tweets were retweeted. The tweet with the highest favourite count also had the highest retweet count. However, the data shows that most tweets favourites did not go beyond the mark 60000.



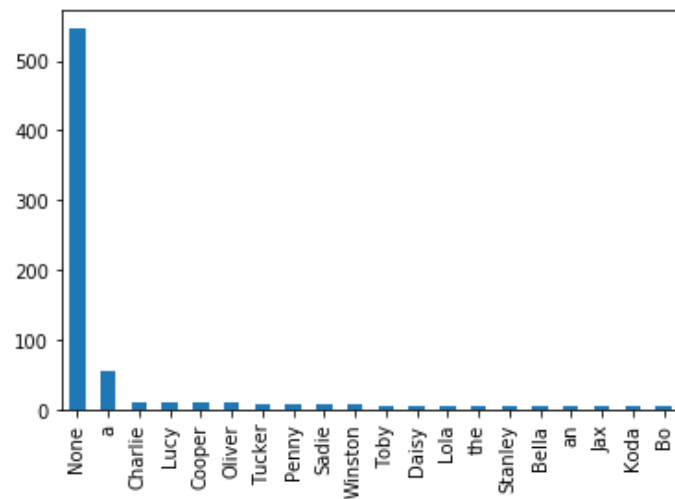
This plot shows how the retweets stood against the rating numerator. The plot shows outliers. The highest rated ones were mostly not retweeted. Most retweets received ratings under quarter of 250.



This plot shows the frequency of retweets. However, one can notice that there is missing data just before and after the 50000 retweets. This will require further investigation into the missing data.



The plots here show the top 20 most predicted dog breed for each prediction. The top 3 breeds are Labrador Retriever, Golden Retriever, and Chihuahua amongst the predictions. According to a [website](#), in 2017 Labrador Retriever was the number 1 breed for most popular breed list, while Golden Retriever sat at number 3. Chihuahua was at number 32.



The plot shows the top 20 most common dog names. Amongst the most common dog names, 'None' and 'a' stand out. However, these names could be disregarded as dummies as there is a probability that these aren't dog names. If one looks at the names of dogs, Charlie, Lucy, Copper and Oliver tend to take the mark.