ANALYSIS DOCUMENTATION

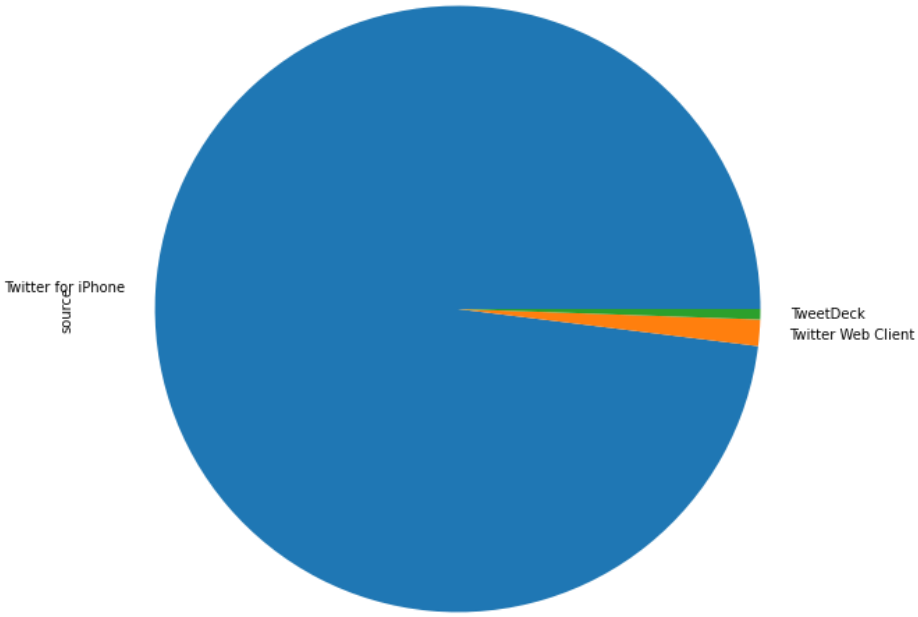
BY: TOLULOPE OLUWADARE

Introduction

WeRateDogs is a Twitter account dedicated to posting photos of Dogs and rating them. The data was provided in three different data types and was cleaned and merged. Among the features pulled together from the three different datasets, the source of tweets and the predictions made on the Dog breeds were of interest to me.

Most Tweet Source

An overwhelming majority of the tweets are from Twitter for iPhone indicating that the tweets were mostly made from an iPhone device. Out of 1994 tweets in the combined data, 1955 were from Twitter for iPhone. This represents nearly 98% of the total Tweets.

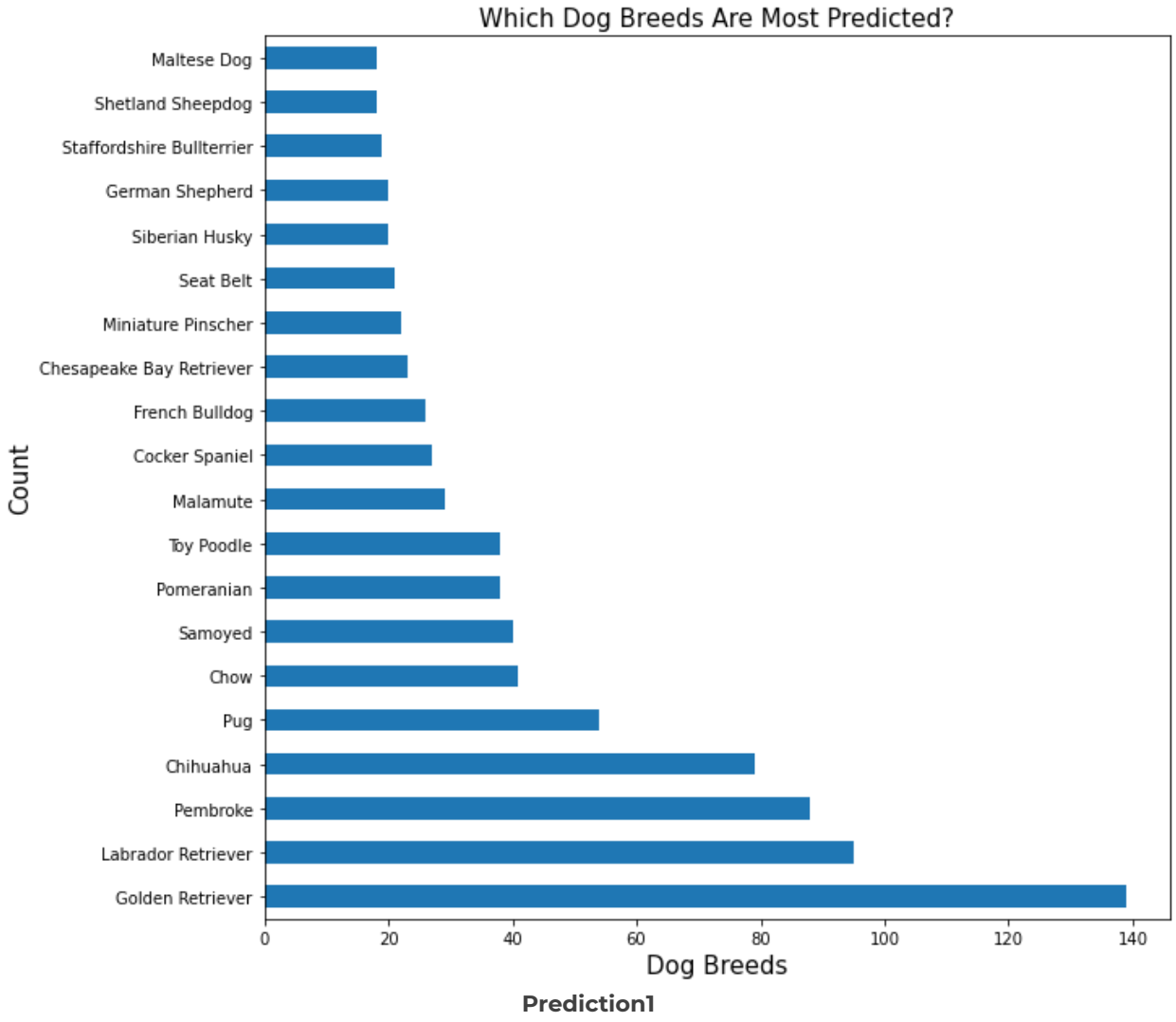


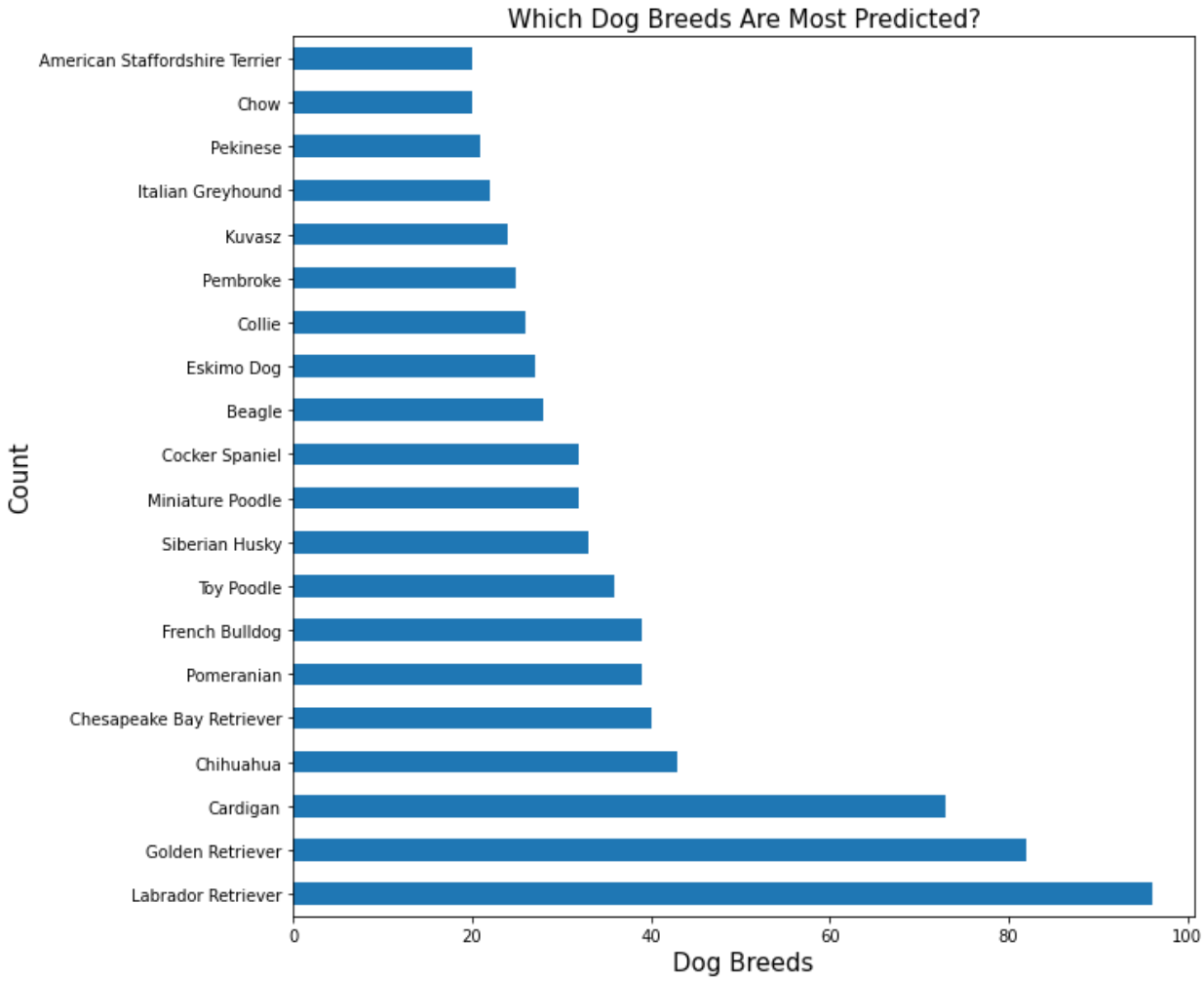
Dog Breed Predictions

For each dog, three predictions were made with varying confidence levels. Each prediction column represents a predicted dog breed and the corresponding confidence value column points to the degree of confidence the algorithm has in that prediction.

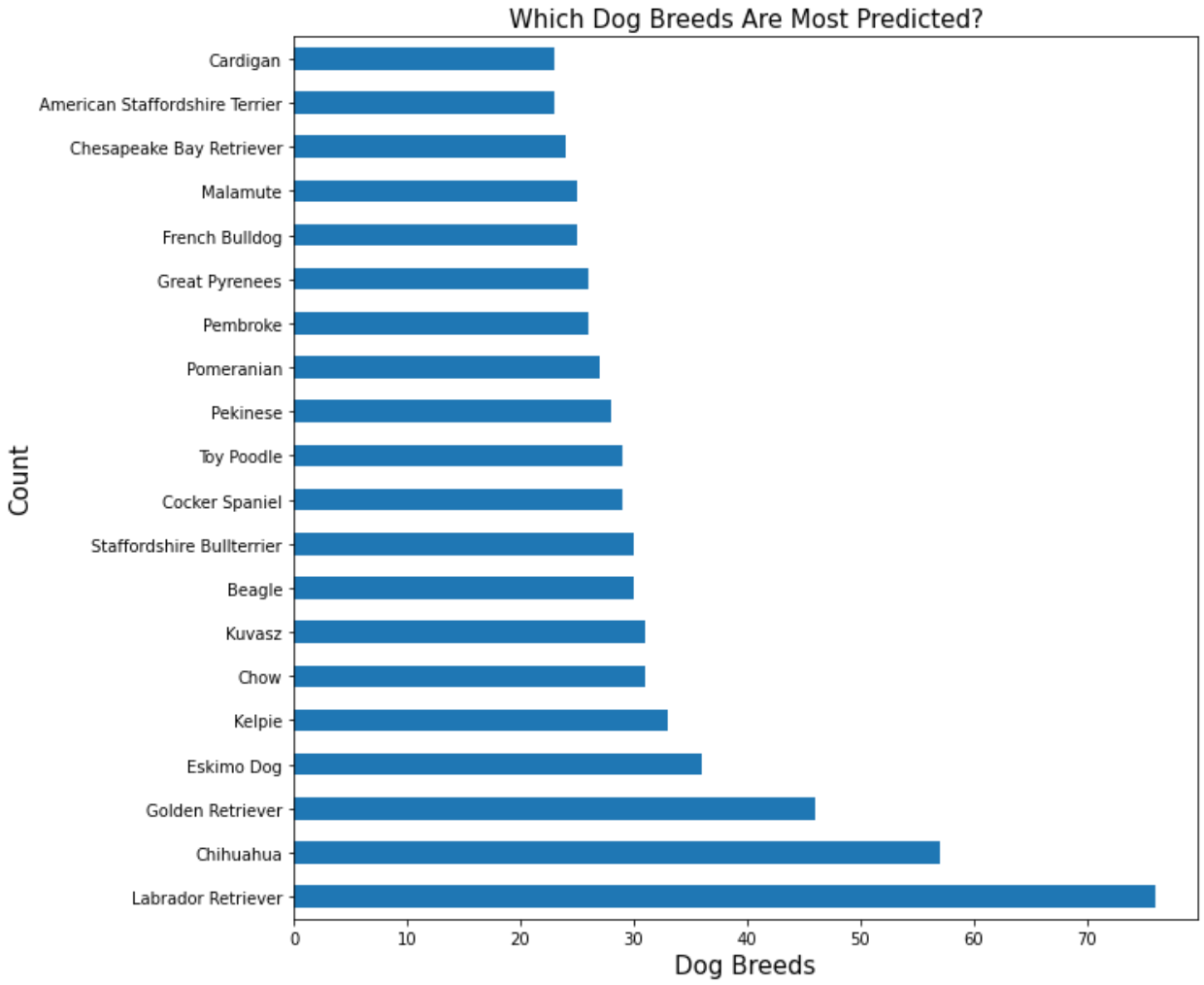
While assessing the prediction columns for the dogs that were mostly predicted, the results were similar yet divergent. The second and third predictions had more similarities in the dog breeds that were most predicted. However, they both scored poorly when the mean confidence percentages were checked. The first prediction had a mean value of 59% while the second and third predictions had 13% and 6% respectively. We can then infer that the column that would have the highest degree of accuracy with respect to the dog breeds is the first column.

It is, therefore, no surprise that the dog breeds that appeared top of the list as most predicted breeds in the first prediction ranked lower or didn't even appear in the second and third predictions. The charts below show the trends in the three predictions:





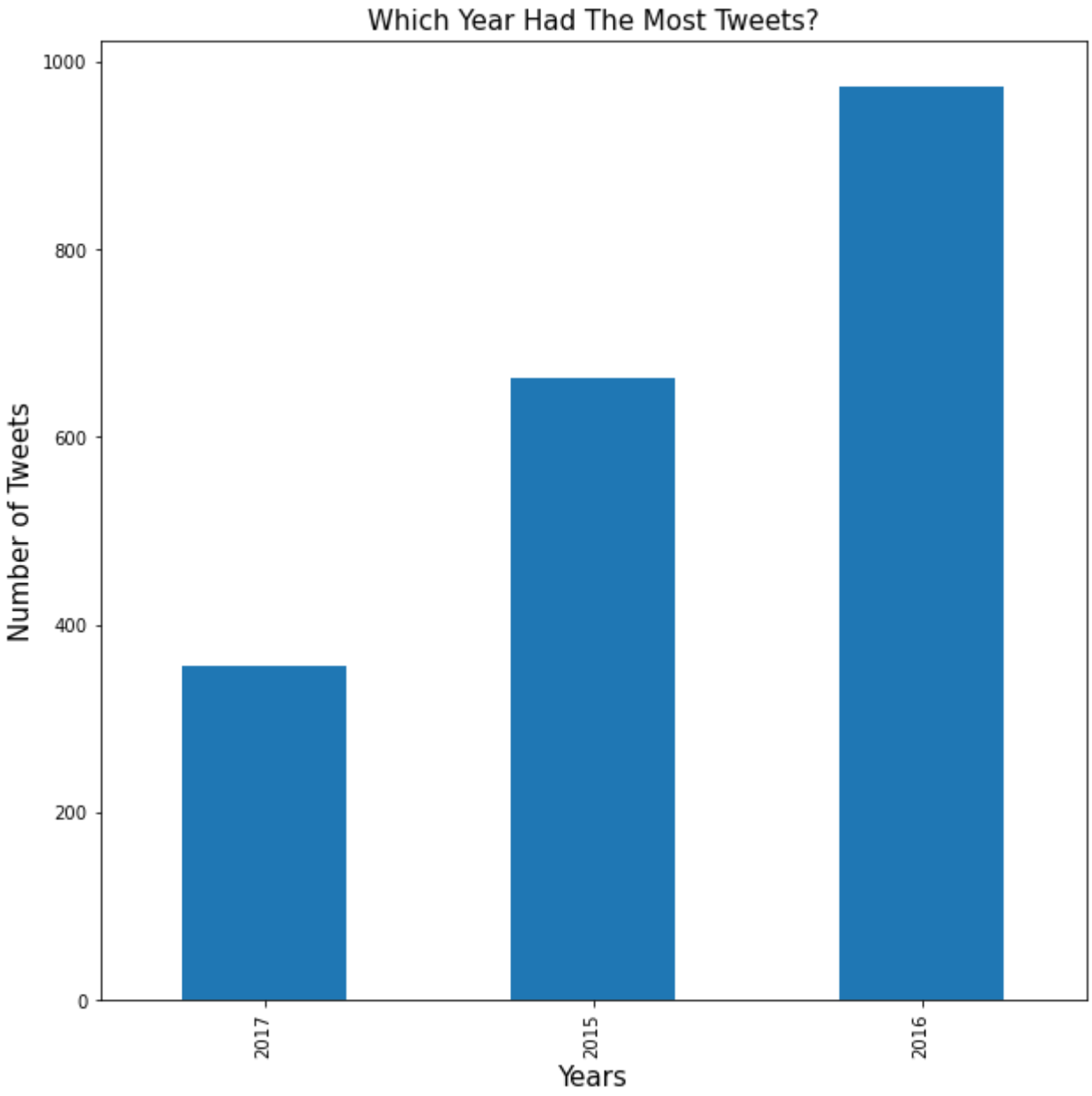
Prediction2

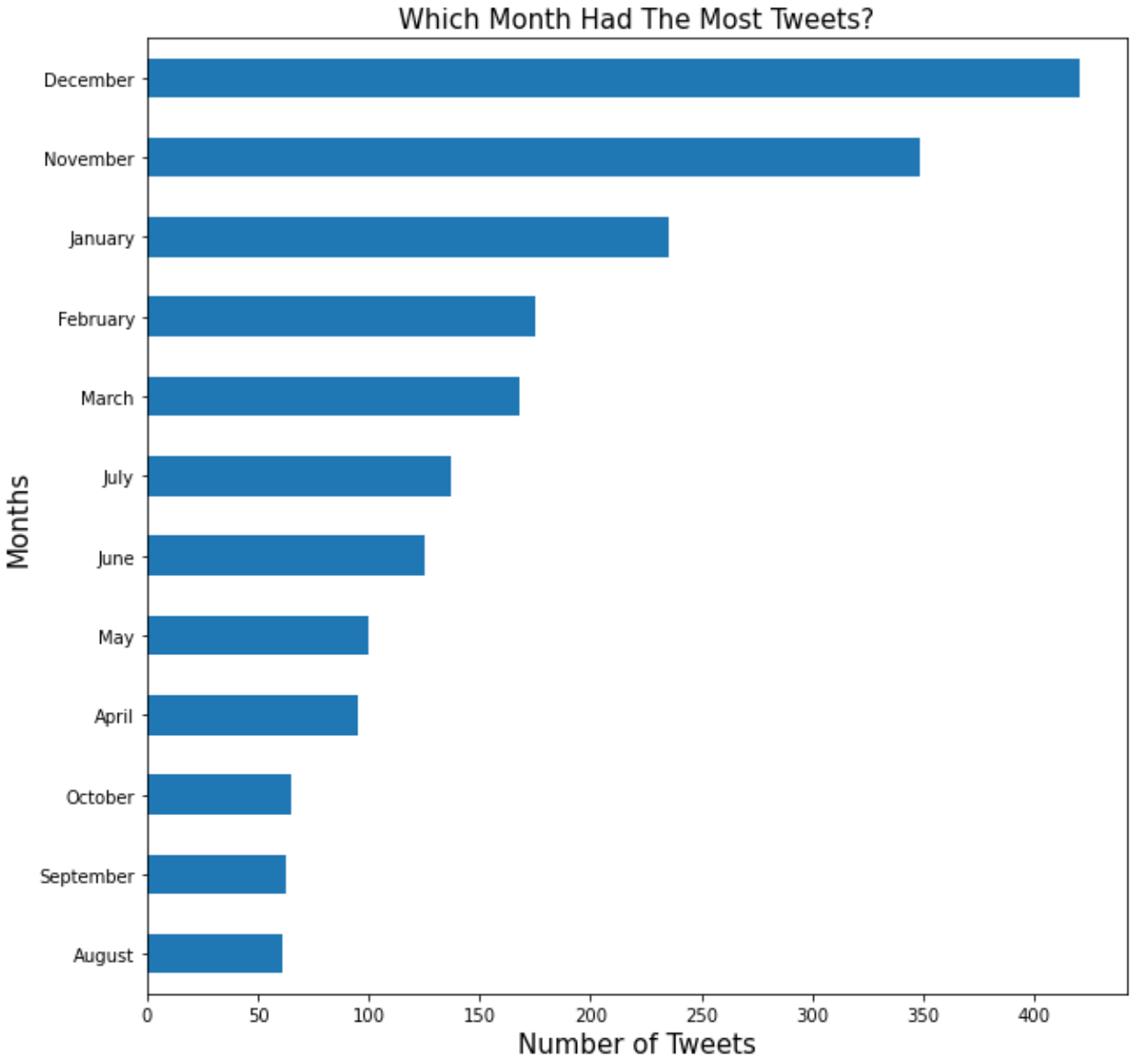


Prediction3

Period of Tweet

An assessment of when tweets were made, provided insightful information on when the WeRateDogs account made the most tweets. On a yearly basis, 2016 had the highest number of tweets with almost 1000 tweets followed by 2015. The dataset held data for only 3 years; 2015, 2016 and 2017. On a monthly basis, the most tweets were made in December. This was followed closely by November, with the first three months of the year trailing after it.




Which Month Had The Most Tweets?
December
November
January
February
March
July
June
May
April
October
September
August
Months
0
50
100
150
200
250
300
350
400
Number of Tweets