

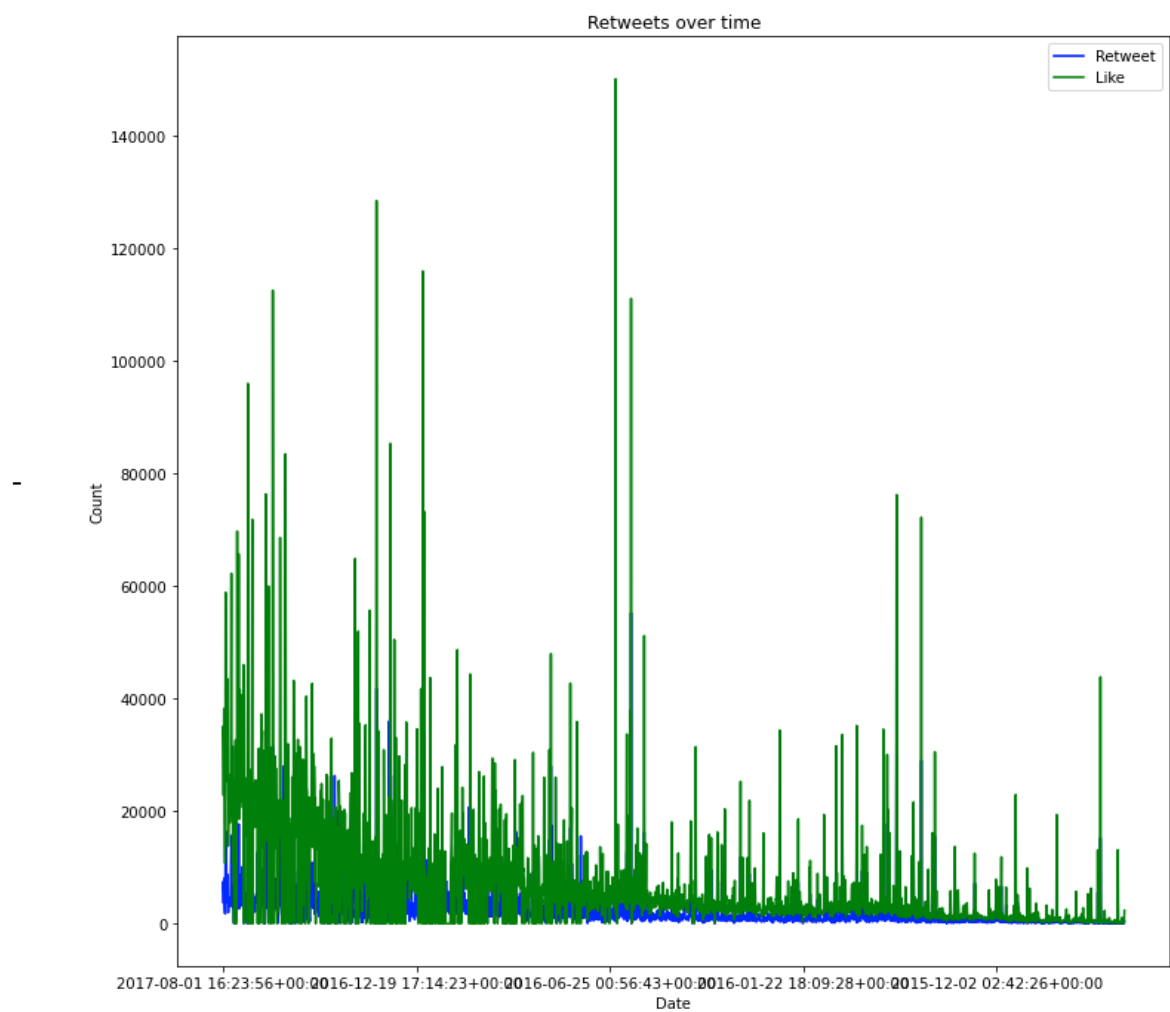
Introduction

The data from WeRateDogs twitter account has been used for the analysis. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. After data wrangling and cleaning processes described in the previous report file (wrangling_report.pdf) data visualizations that explain the data findings have been created in Python using matplotlib library.

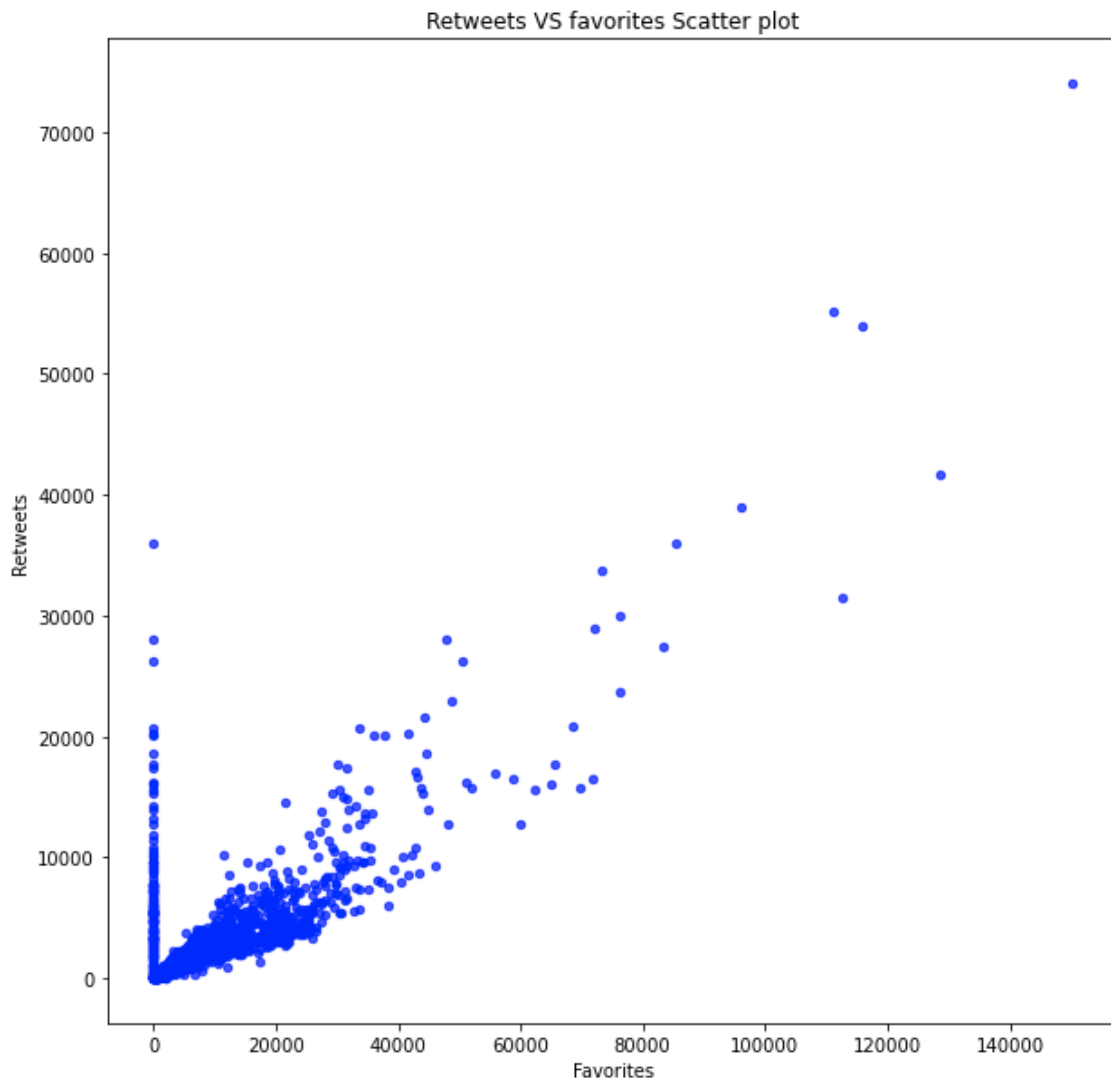
Using different charts -combo column chart, scatter plot and box plot the trend of profile popularity over time ,the correlation between retweets and likes and which is the dog category with most participants and which dog category is with highest retweets.

Analysis

- 1) For the creation of the below chart the “retweet_count” and “favorite_count” columns have been used. Based on the created combo column chart below we can assume that the number of likes(favorite_count) are a lot higher than retweets. The trend of retweets and favorites are increasing over time. The twitter profile becomes more popular over time. Moreover the frequency of high counts from retweets and favorites appear more and more. (The chart describes the time backwards- from end to beginning).



- 2) The scatter plot best describes the correlation between retweets and favorites. From the chart we can assume that the relation between retweets and favorite counts are proportional that means with the increase in the number of likes the retweets also increase.



- 3) The dogs in the site are classified in four different stages- doggo, puppo, floofer and pupper. To describe and visualize the dogs category I have chosen to use a box plot. From the created visualization I can deduce that most dogs are in the Puppo dog category but highest retweets are always about doggo dog category.

