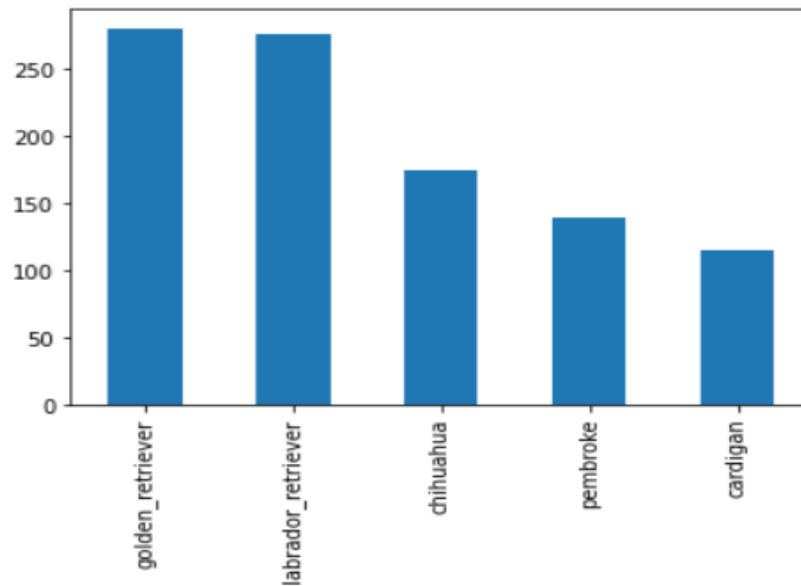# Act Report

1. Storing Data

- I have merged all the 3 data frames into one data frame called master_data and I have saved it in twitter_archive_master.csv file
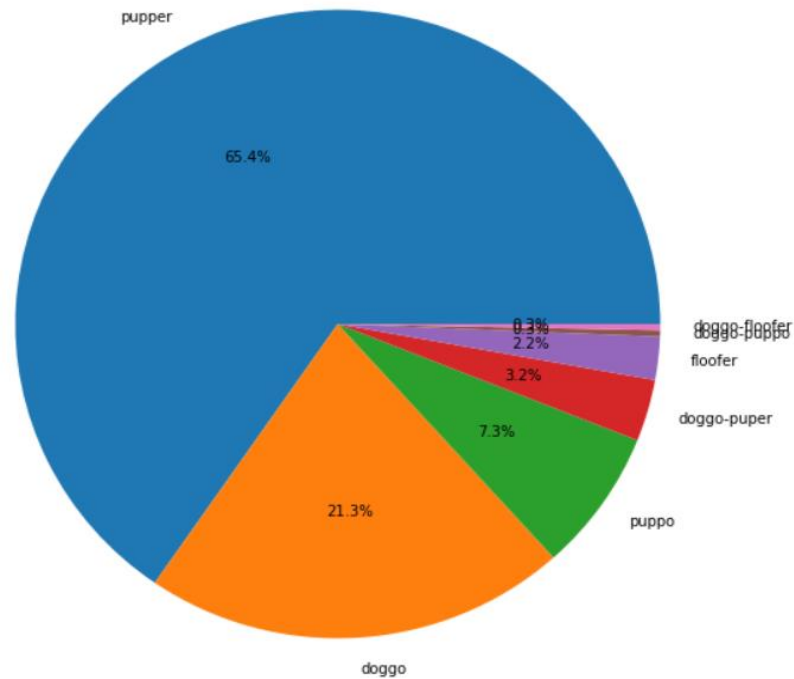
2. Analyzing and visualizing data

### 1. Bar chart for the most famous five dog breeds

- As shown in the graph it seems that the golden_retriever and Labrador_retriever breeds are the most widely famous breed of dogs with score more than 250.

## 2. Pie chart shows the different percentages for each dog stage

- The Pupper stage percentage: 65.4%
- The Doggo stage percentage: 21.3%
- The Puppo stage percentage: 7.3%
- The Doggo-pupper stage percentage : 3.2%
- The Floofer stage percentage : 2.2%
- The Doggo-puppo stage percentage : .3%
- The Doggo-floofer stage percentage: .3%

- As shown in the figure it seems that the pupper stage has the highest percentage equals to 65.4 % and both of doggo-puppo and doggo-floofer have the lowest percentage equals to .3 %

## 3. Bar chart shows the proportion between the retweet counts and favorite counts for each dog stage

Proportion by favorite couts, retweet_counts and dog_stage
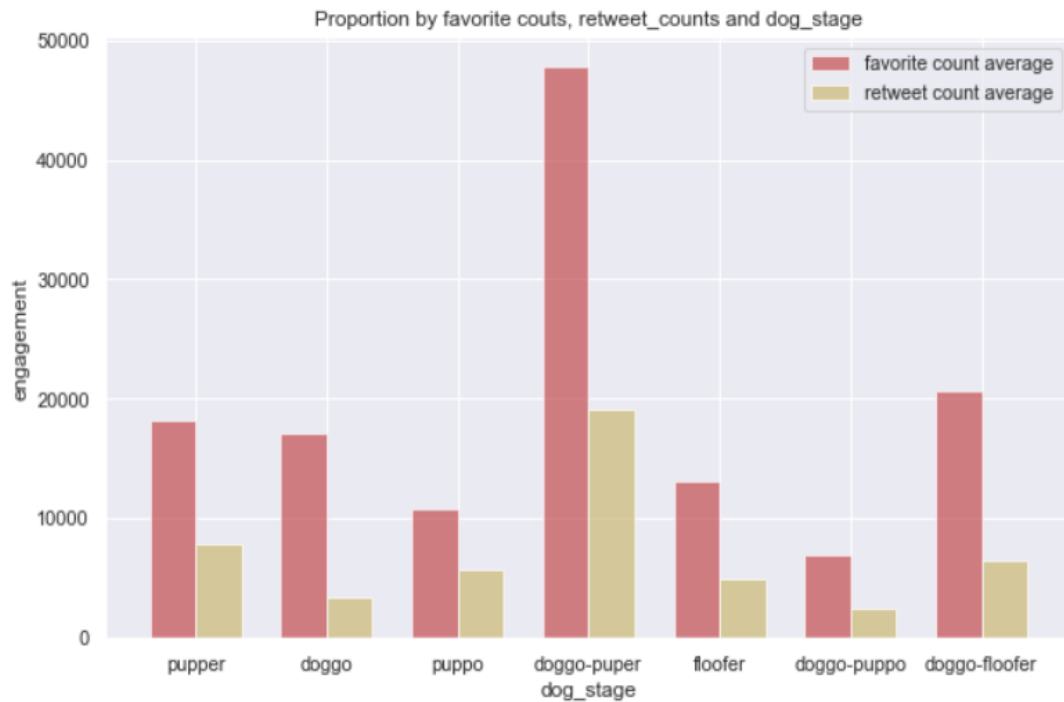
- As shown in the graph it seems that the doggo-pupper dog stage got high favorite count average more than its retweet count average.

4. Scatter plot shows the relationship between the favorite counts and retweet counts
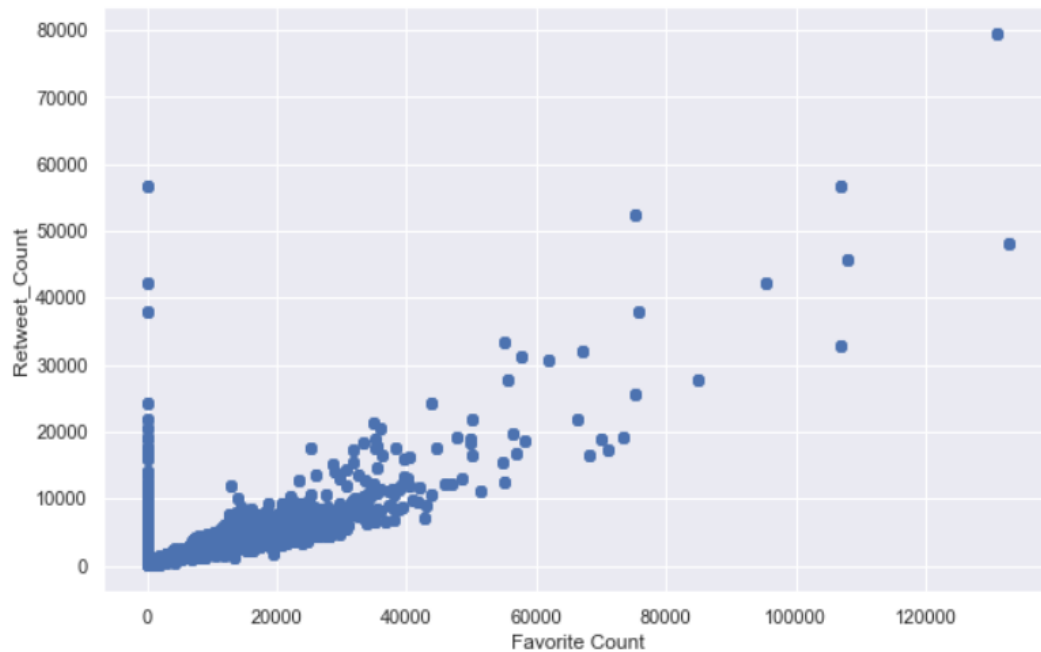
- The relation between favorite counts and retweet counts is positive and the distribution between them seems to be right skewed.

5. timeseries plot shows how much rating numerator changes over time
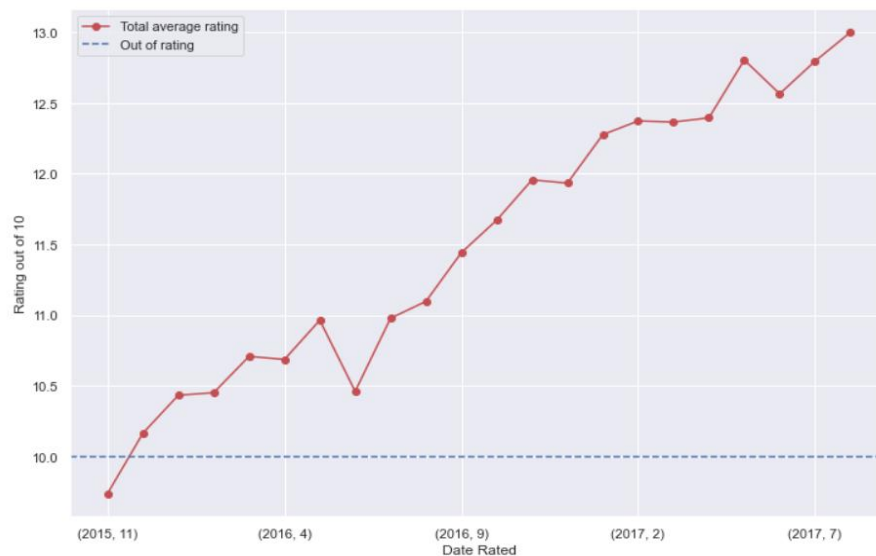
- As shown in the graph the rating numerator increases over the time