

# Act Report

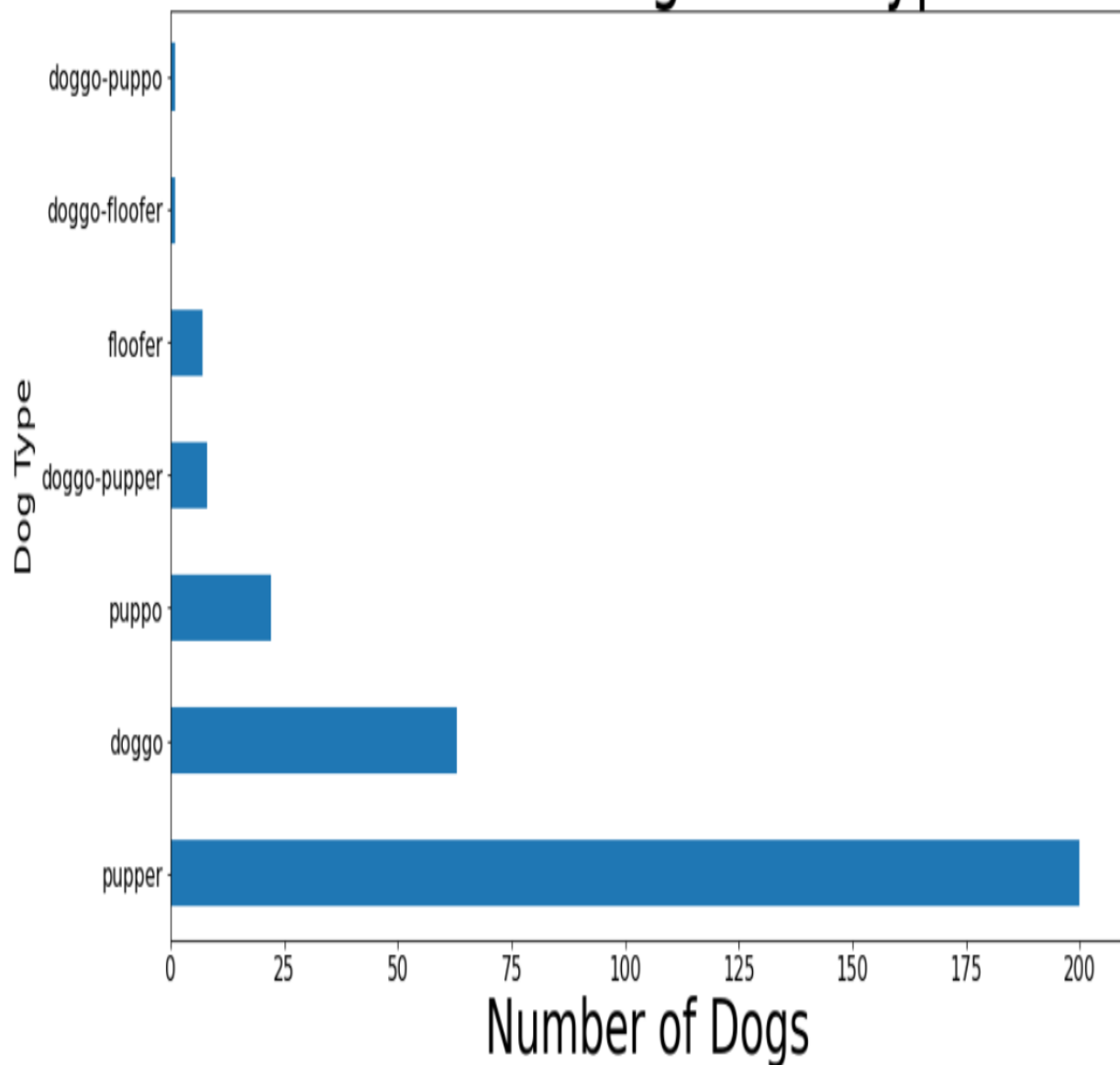
- Real-world data rarely comes clean. Using Python and its libraries, I gathered data from a variety of sources and in a variety of formats, assessed its quality and tidiness, then cleaned it. This is called data wrangling.
- The dataset that I wrangled (and analyzed and visualized ) is the tweet archive of Twitter user **@dog\_rates**, also known as **WeRateDogs**. WeRateDogs is a Twitter account that rates people's dogs with a humorous comment about the dog. These ratings almost always have a denominator of 10. The numerators, though? Almost always greater than 10. 11/10, 12/10, 13/10, etc. Why? Because "**they're good dogs Brent.**" WeRateDogs has over 4 million followers and has received international media coverage.

## The analysis and Visualizations

### The Insights

1. What is the famous dog type between dogs ?
  - a. By looking to the data I saw that people called a name or type about the dogs and they are 4 categories ( floofer – puppo – doggo – pupper ).
  - b. In this visualization I used matplotlib python library to plot.
  - c. The x-axis is **Dog type** and y-axis is **Number of dogs**.
  - d. By looking to the graph we can see that most famous type or category of dogs is “pupper” 😊

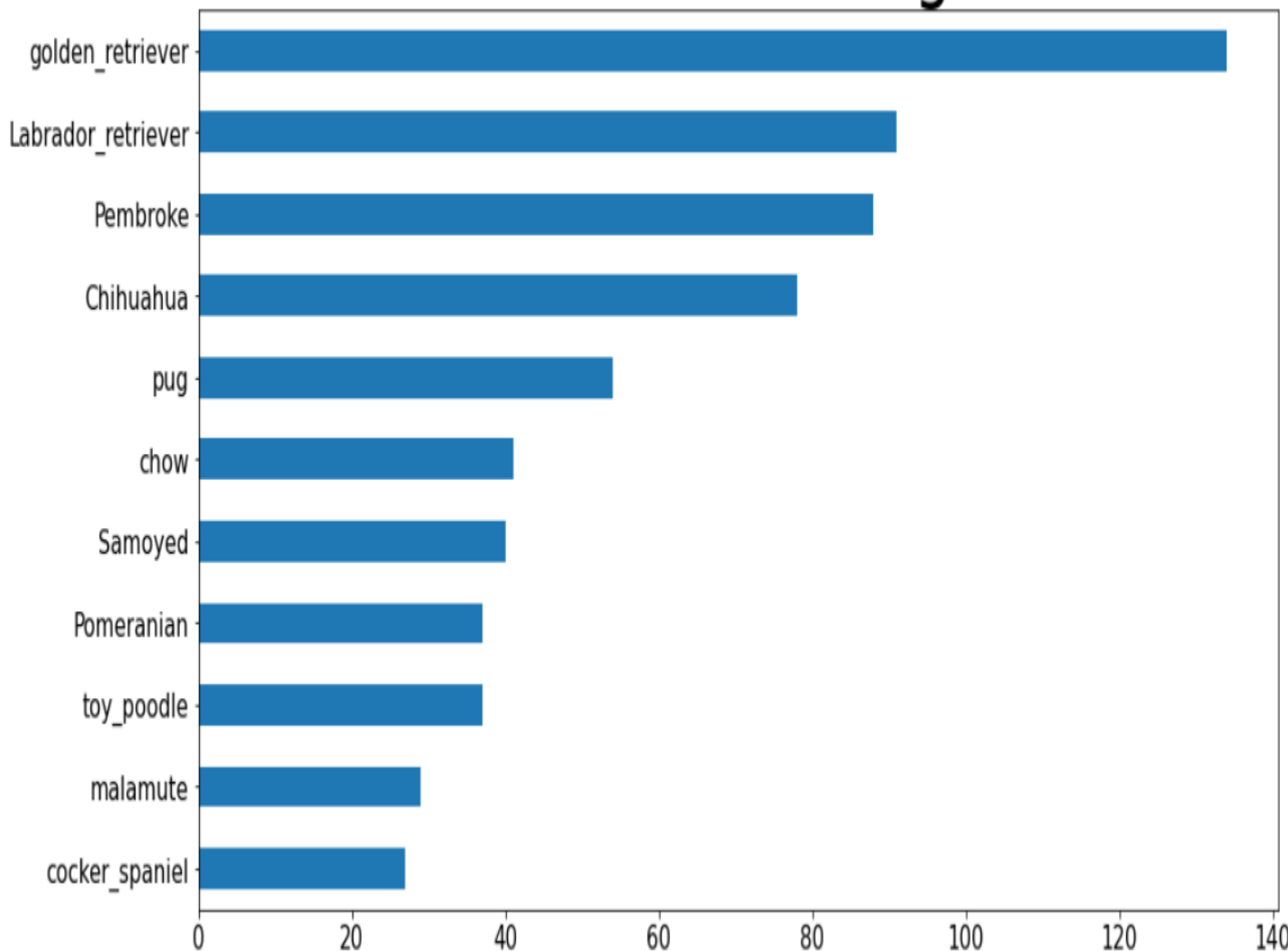
# Number of Dogs for 4 Types



2. What is the most famous breed between fans ?
  - a. By looking to the image prediction dataframe I realized that there is a neural network which helps in the prediction of content of the images which there urls also within the data.
  - b. Here also I used matplotlib python library to plot the graph.

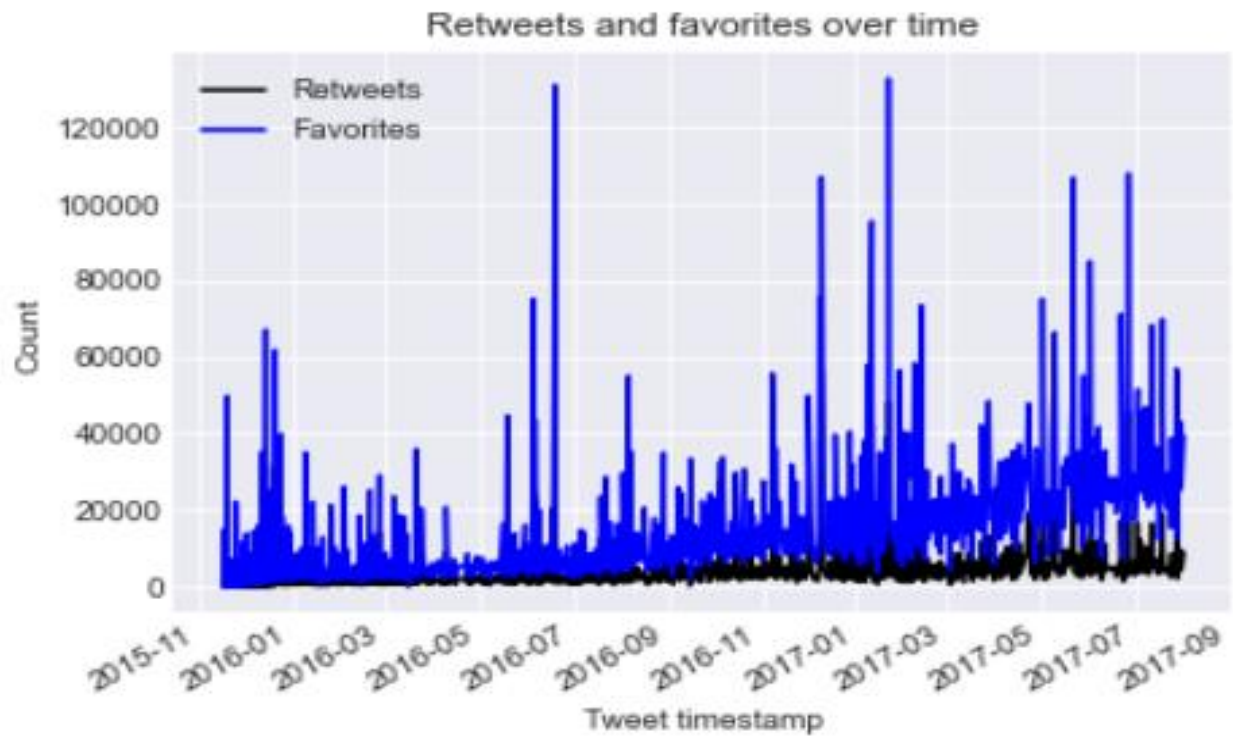
- c. In the Y-axis the dog breeds and x-axis is the number of dogs of each breed.
- d. It's found that most common dog breed is golden retriever.

## 10 Most Common Dog Breeds



- 3. Distribution of retweet and favorite counts over time.
  - a. In the Y-axis the Count and x-axis Tweet timestamp.
  - b. Here also I used matplotlib and seaborn python libraries.

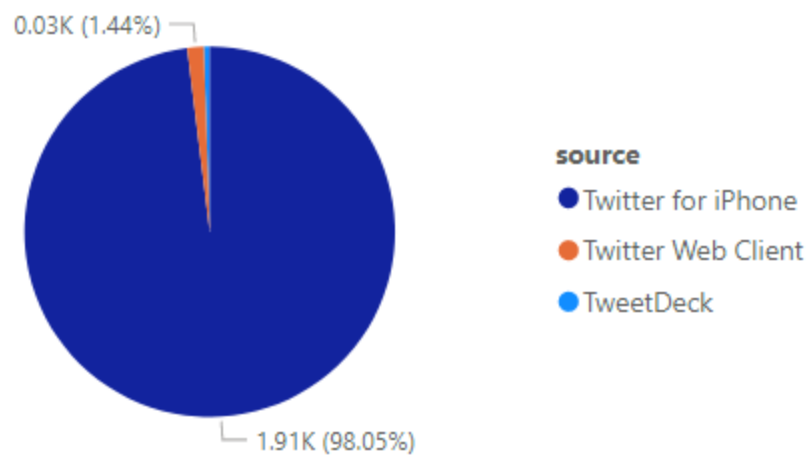
- c. We can see here that there is increase in retweets and favorite counts over time.



## The Visualizations

1. Percentage of Each source of data

## % Of Each Source Of Data



## 2. Number of tweets per person

## Number Of Tweets Per Person

name ● Charlie ● Cooper ● Oliver ● Lucy ● Tucker ● Penny ● Winston ● Sadie ● Daisy ● Toby ● Lola

