

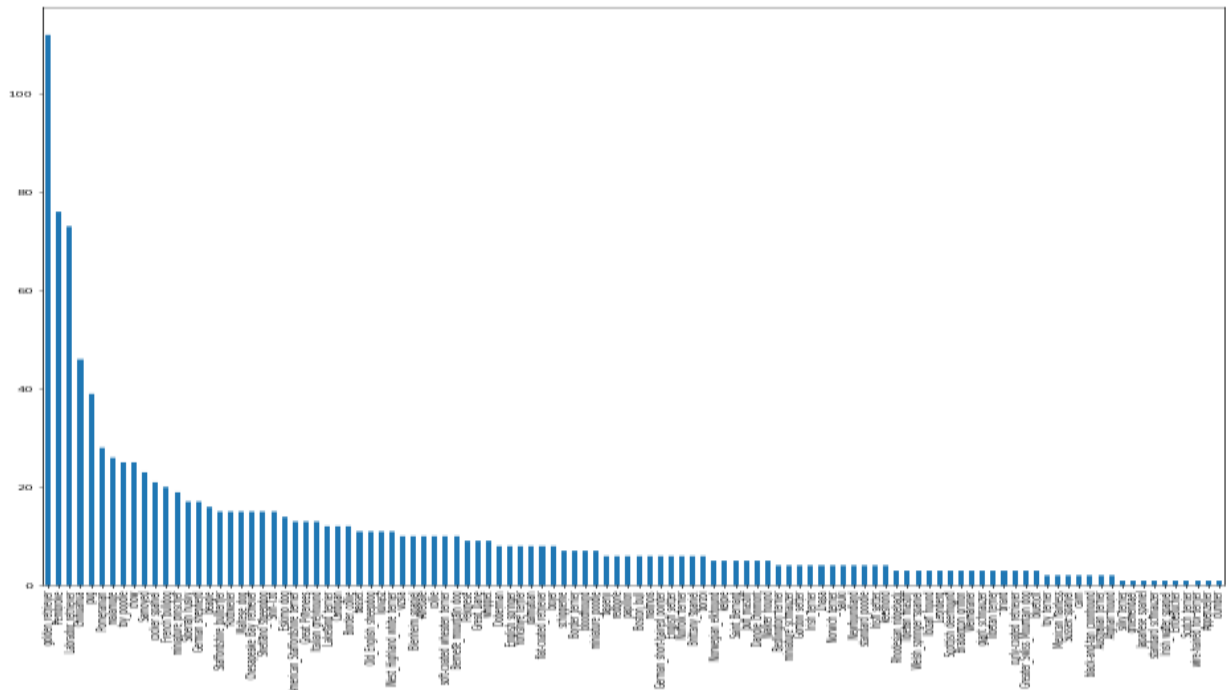
VISUALIZATIONS & INSIGHTS

In this report we are about to see each insight and visualization I included in my project, and what questions they exactly satisfy, I tried to find some interesting question to answer as well as some cool insights to share.

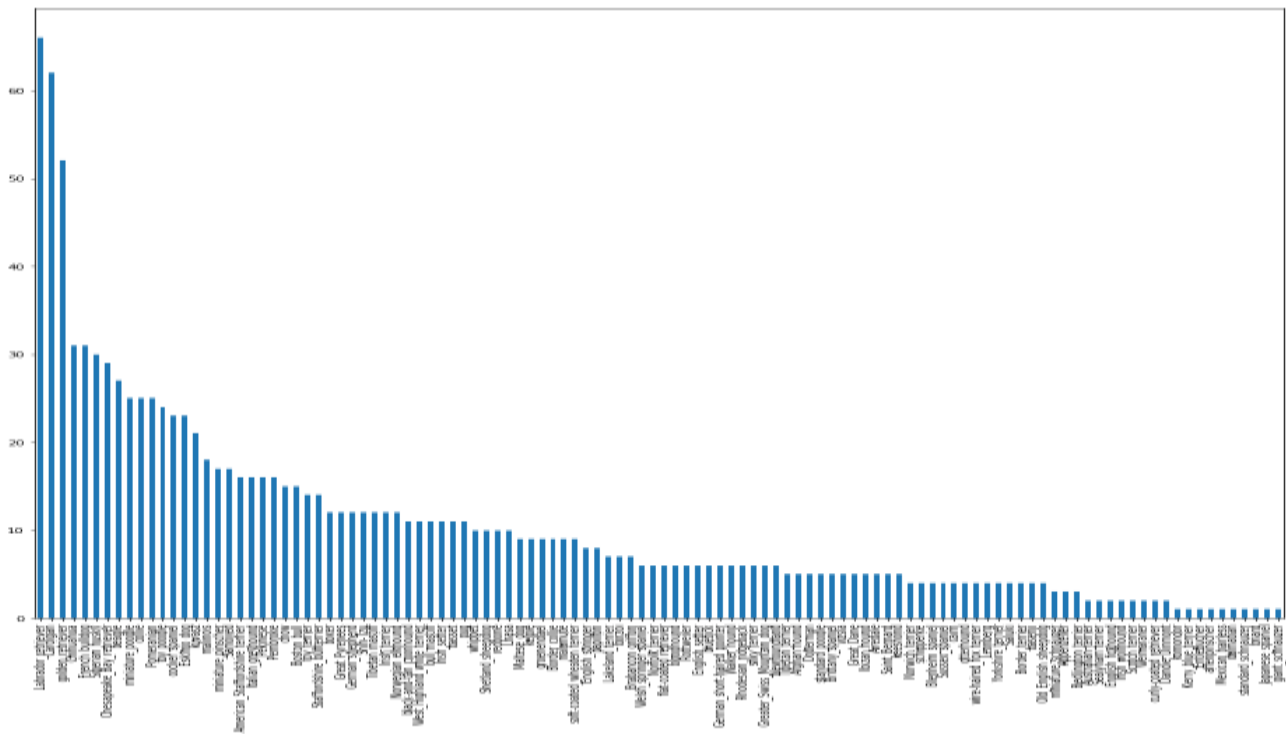
1. What is the most popular dog that is found in our gathered data?

For this question I used the columns of predictions algorithm the first, second and third predictions. I used bar chart for each one, and deduced that the golden retriever is the most popular dog based on the charts

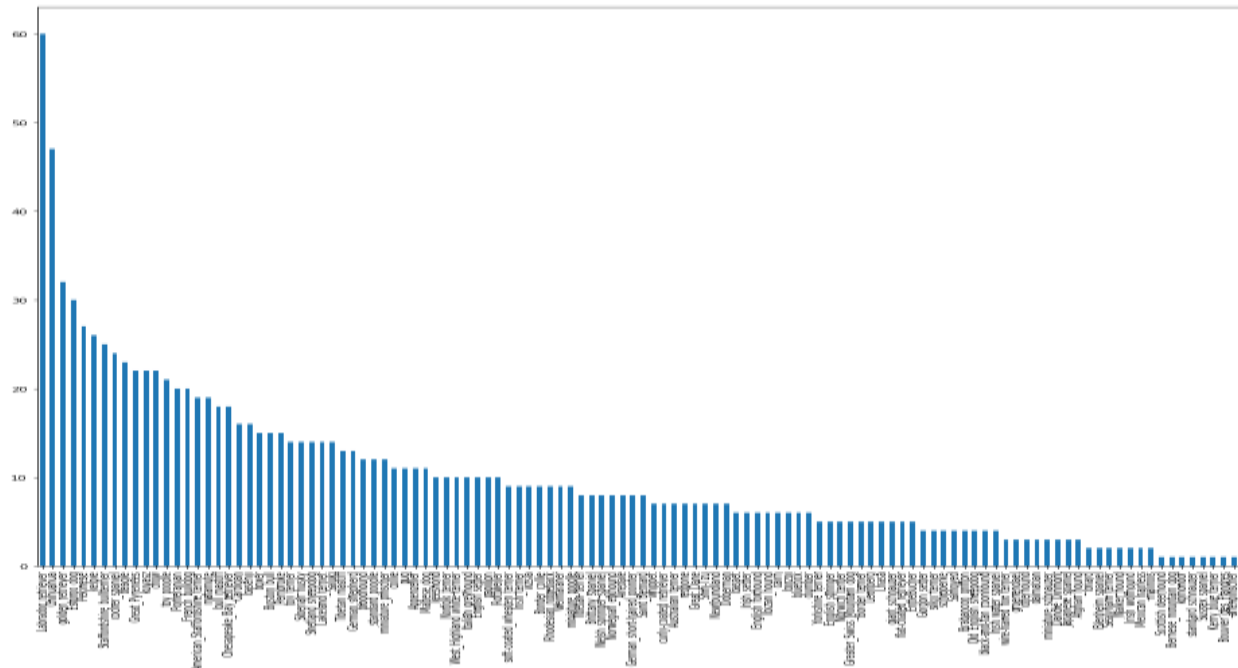
First prediction:



second prediction:



Third prediction:



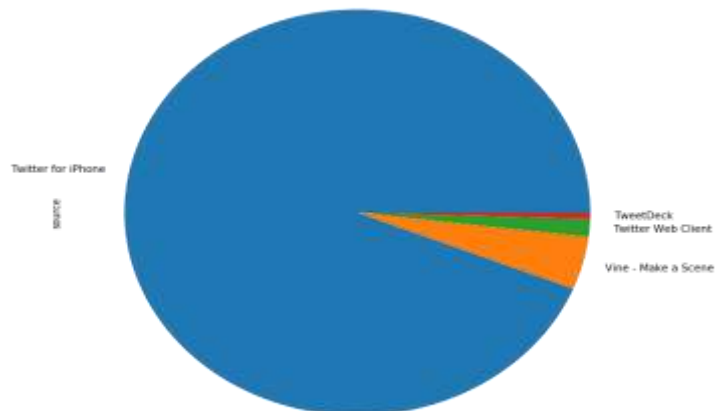
From those charts I found that the most popular dog types are:

- 1- Golden retriever
- 2- Labrador retriever
- 3- Chihuahua

2. The most used source for the tweets:

For that insight, I used the source column from the merged and cleaned data set and visualized the counts by using a pie chart:

Form the chart I found that the most used source is Twitter for iPhone.



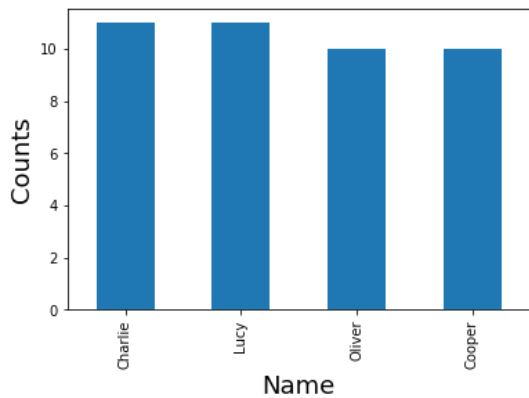
3. The second most favored dog:

In this one I used the favorite_count to know the maximum count and by the index I got the image of the second most favorite since the first doesn't have an image URL.



4. The most used name is ...

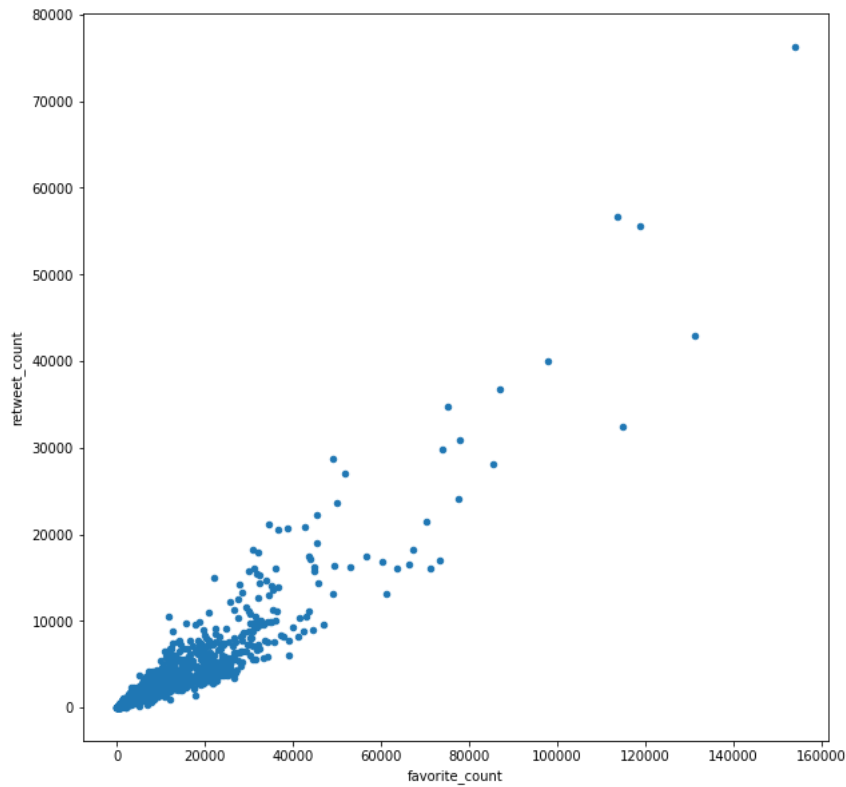
Here I used the name column to get the counts for each name, and then visualized the most popular 4 names used:



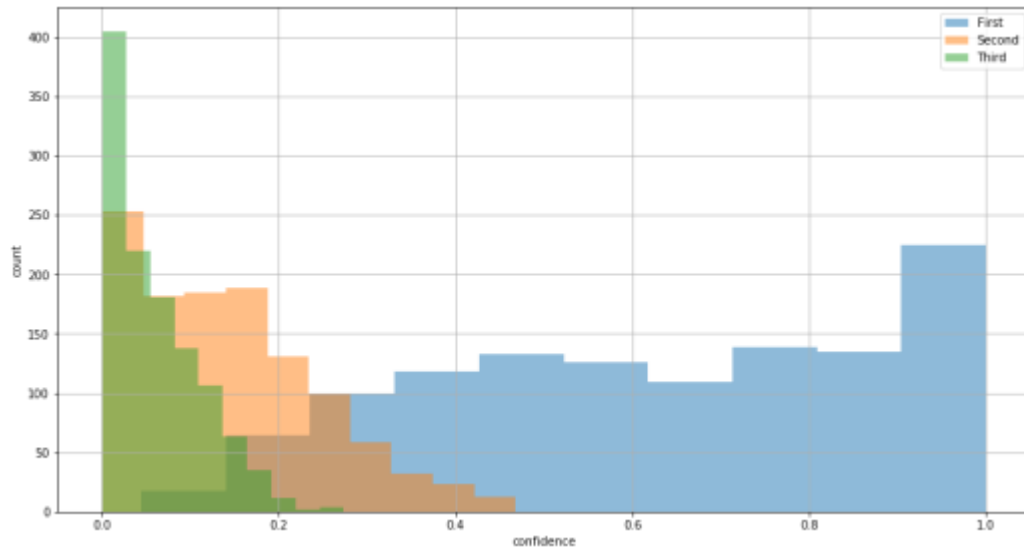
And it was 'Charlie'.

5. Positive correlation between favorite_count and retweet_count

By plotting a scatter plot between those two columns I found that they have a positive correlation (only natural I guess :D)



6. The highest prediction confidence is for the first algorithm used



I was curious about which prediction algorithm of the three has the maximum confidence result and accordingly has the most correct prediction, for that I plotted the three histograms of prediction confidences together in the same plot and it is so obvious that the first prediction is the winner.

7. [The most retweeted tweet](#)

For that one I used the retweet_count column to get the maximum count of retweet and then it is tweet URL and opened it in the browser.