Insights and Visualizations
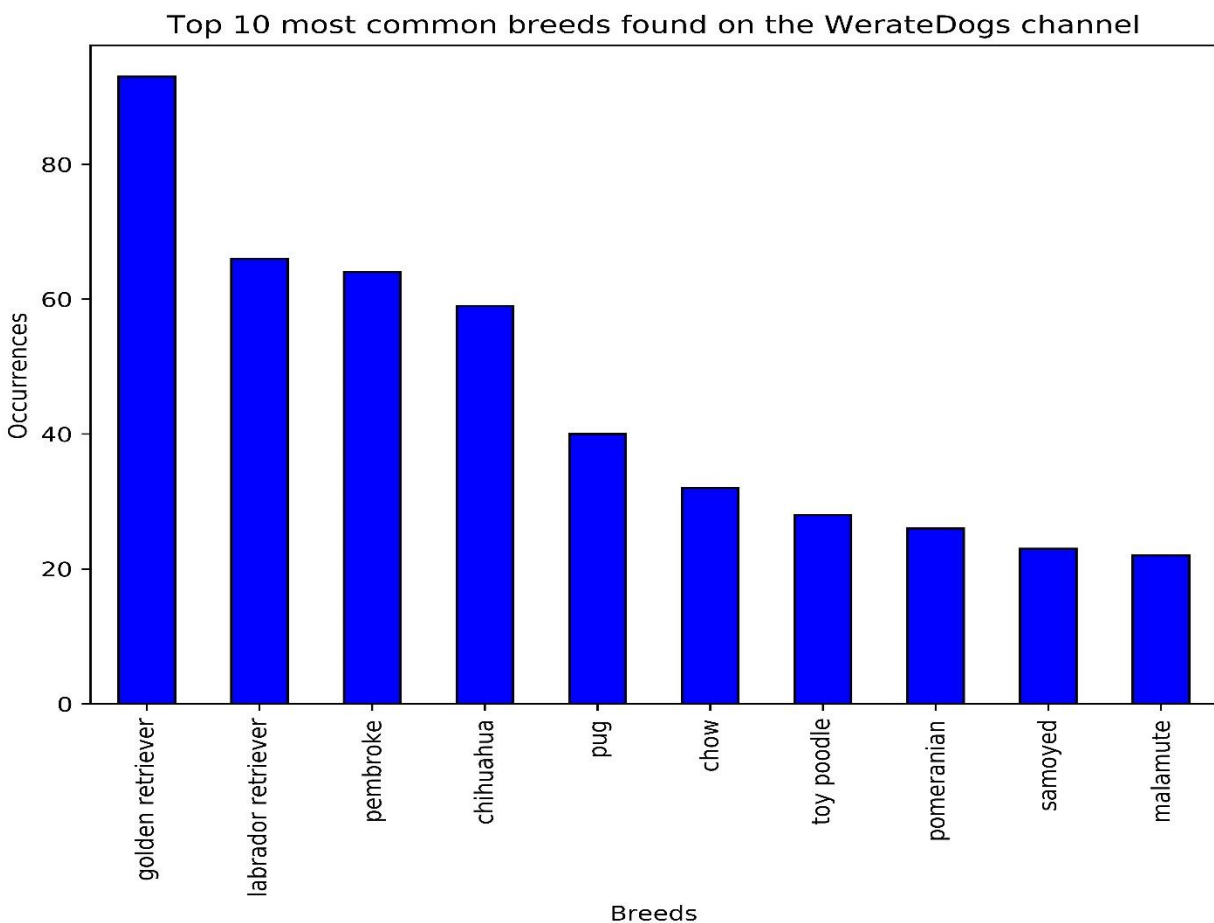
I had some hands-on experience on the data analysis process such as:
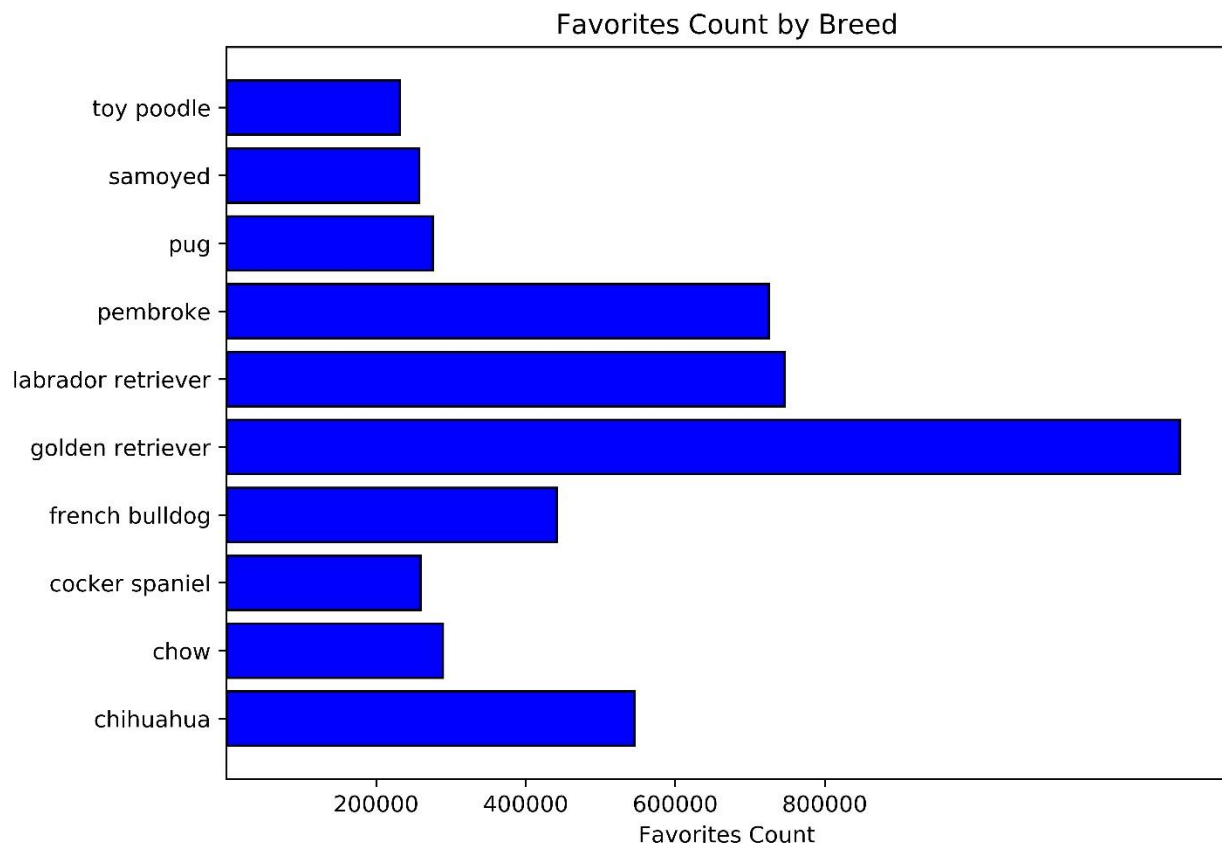
1) With gathering data (using tweepy module)
2) With Data wrangling of the gathered dataset .
3) With Cleaning the dataset after assessing it.
4) With Visualizing it and putting forth insights accordingly.

All In all this project has given me the required confidence to tackle any data analysis problem in the future.
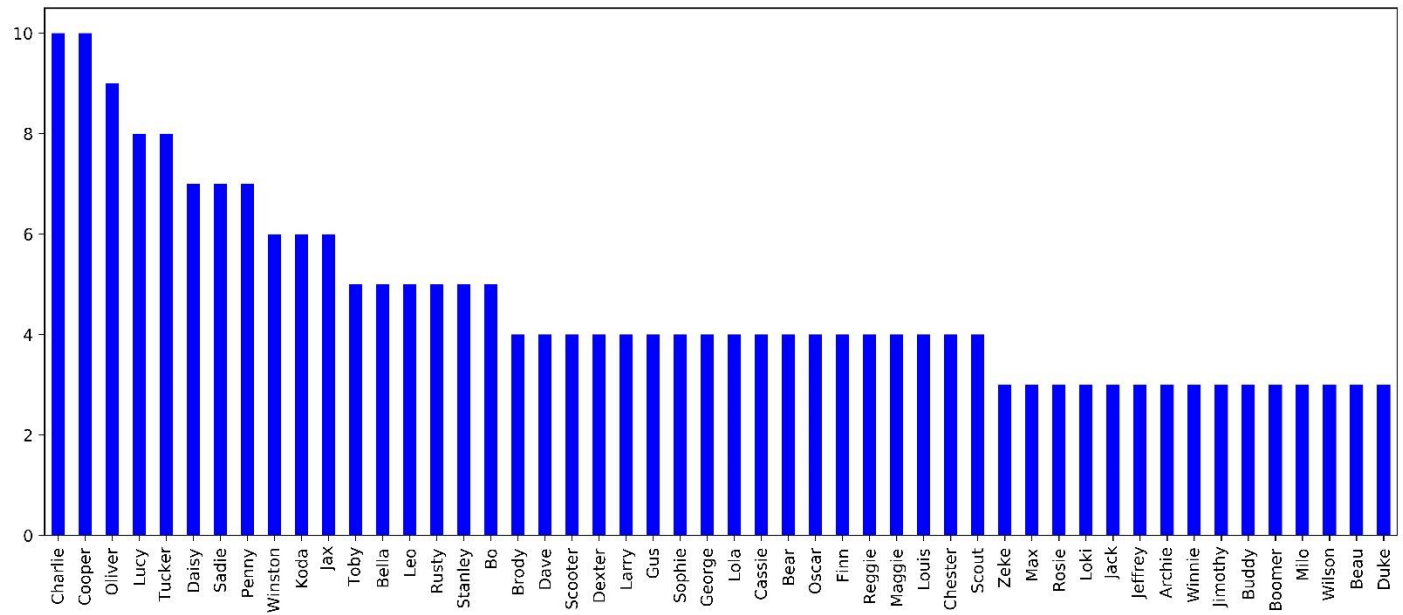
I will surely be using the template provided in the wrangling process as it has helped me in developing the good habit of writing good notes , and also making it a step in the wrangling process makes my work more organized.

Top 10 most common breeds found on the WerateDogs channel

We can deduce from the above visualization that the golden retriever and Labrador retriever are the most common breeds found in our dataset.

Favorites Count by Breed

After grouping the table by predicted breeds and them summing the favorite count for each of the breeds, I got this list of top 10 most favorited dog breeds from the weratedogs channel dataset. No surprise again that the golden retriever leads the favorite dog breed by more than 200000. Second in line is Labrador retriever followed closely by Pembroke.

Finally, I thought of looking at the most common names for dogs nowadays. Calculating that was simple enough I just grouped the dataset by name column and plotted the sum counts of occurrences of each name, but only the top 50 were considered. Turns out Charlie and Cooper are the top choices of people according to the weratedogs dataset with each having 10 unique instances.