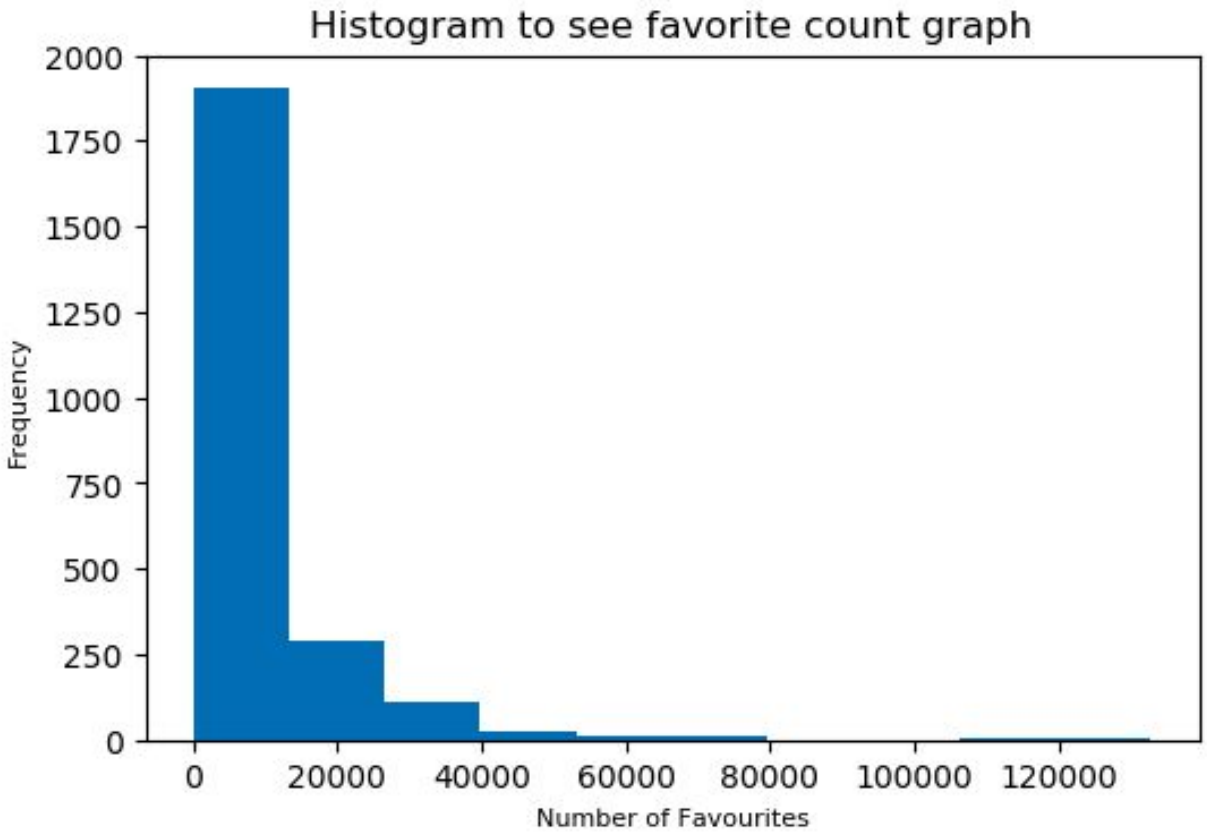
# Data Wrangling: We Rate Dogs

I analysed a Dataset provided by Udacity on Dogs rating rated by We Rate Dogs ( A Famous Twitter account having more then 4 millions followers, they rate everyone's dog around the globe). Please like/ retweet and reply to those tweets and show their emotions.  Here are some Useful insight's which I draw by analyzing this data provided by Udacity.

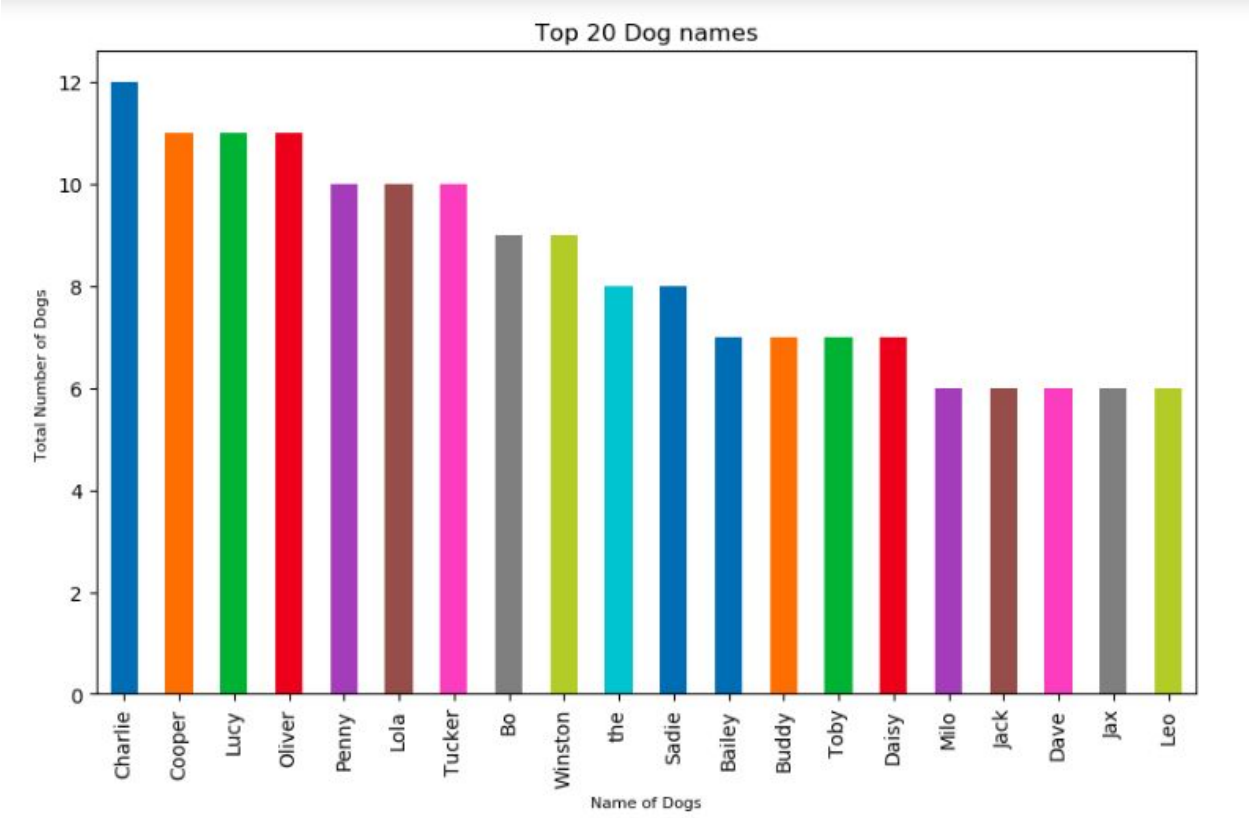
**Q**: Which is the most favourite dog among these tweet's?
**A**: The most favourite dog is from 'Puppo' Dog type, having favorite_count = **132810** on a single tweet by different sorts of people. Here is the Tweeted link we get from the data: https://twitter.com/dog_rates/status/822872901745569 73.

**Q**: That was the Ratio of Favourites every tweet got.
**A:** The Average Likes/Favorites a normal post got were **8080.** I also Draw a Histogram to see the flow of favourite count.



**It can be easily seen that the most of the tweets are under the digit on 10,000 favourite count. The Major portion of tweet's got the same favouritism. It is crystal clear by this analysis.**

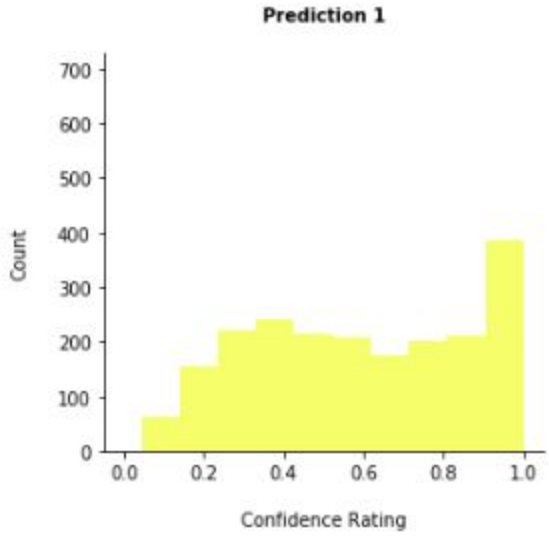
**Q:** The Most Popular Top 20 Dog Names.
**A**: During the Analysis I've seen that the most popular dog name was Charlie and it was chosen by people for their dogs most of the time.
Please have a look towards the whole graph with Visualization to see how many people like to call their dogs with these top 20 names.

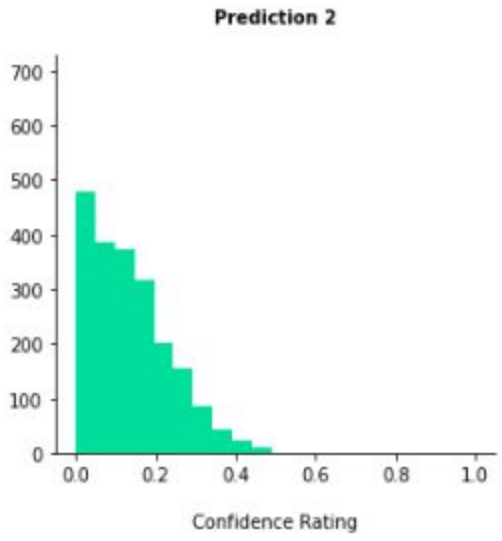


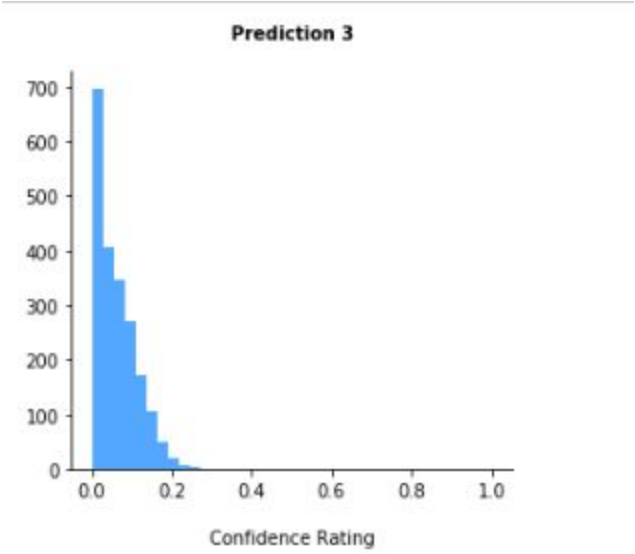

The Bar Chart shows the Does name with different color combinations. The Number of total dog names are on Y-axes where the Name is at the Bottom of the graph. ( on X-axis).

**Q**: Compare the Predictions (via Neural Network) with Actual Tweet's to have comparisons between them. I'll show here two samples.

**A**: First, let me show you the graph for the confidence level for each prediction In the Dataset we have 3 different predictions for each tweet. Let's have a look towards it's visualization.




The number of images are on y-axes and the confidence level are on x-axes. Total Confidence level is 1 and it lies between 0-1. From Prediction and Prediction3, we can say that between 0.0 to 0.2. Now' let's see it in action.

Suppose we have sample of 4 tweet Ids, Details are below.

:

| | tweet_id | jpg_url | img_num | prediction_order | prediction | confidence | dog |
|---|---|---|---|---|---|---|---|
| **2940** | 707315916783140866 | https://pbs.twimg.com/media/CdDkEkHWwAAAeUJ.jpg | 2 | 1 | Bernese_mountain_dog | 0.979235 | True |
| **5655** | 847842811428974592 | https://pbs.twimg.com/media/C8QkidrVYAQXQh7.jpg | 1 | 1 | Bernese_mountain_dog | 0.951337 | True |
| **5295** | 826598365270007810 | https://pbs.twimg.com/media/C3iq0EEXUAAdBYC.jpg | 1 | 1 | French_bulldog | 0.628119 | True |
| **2520** | 694669722378485760 | https://pbs.twimg.com/media/CaP2bS8WYAAsMdx.jpg | 2 | 1 | beaver | 0.457094 | False |
| **450** | 668641109086707712 | https://pbs.twimg.com/media/CUd9ivxWUAAuXSQ.jpg | 1 | 1 | vacuum | 0.432594 | False |

Now, We are checking the image of Dog for the **first** sample.


Now, let's look towards prediction for this dog from our neural Network data.

| | tweet_id | jpg_url | img_num | prediction_order | prediction | confidence | dog |
|---|---|---|---|---|---|---|---|
| **2940** | 707315916783140866 | https://pbs.twimg.com/media/CdDkEkHWwAAAeUJ.jpg | 2 | 1 | Bernese_mountain_dog | 0.979235 | True |
| **2941** | 707315916783140866 | https://pbs.twimg.com/media/CdDkEkHWwAAAeUJ.jpg | 2 | 2 | Shetland_sheepdog | 0.011037 | True |
| **2942** | 707315916783140866 | https://pbs.twimg.com/media/CdDkEkHWwAAAeUJ.jpg | 2 | 3 | Appenzeller | 0.003971 | True |

Another Example:


And the Predictions, We have for this Dog.

| | tweet_id | jpg_url | img_num | prediction_order | prediction | confidence | dog |
|---|---|---|---|---|---|---|---|
| **5295** | 826598365270007810 | https://pbs.twimg.com/media/C3iq0EEXUAAdBYC.jpg | 1 | 1 | French_bulldog | 0.628119 | True |
| **5296** | 826598365270007810 | https://pbs.twimg.com/media/C3iq0EEXUAAdBYC.jpg | 1 | 2 | Siamese_cat | 0.117397 | False |
| **5297** | 826598365270007810 | https://pbs.twimg.com/media/C3iq0EEXUAAdBYC.jpg | 1 | 3 | cougar | 0.082765 | False |

It can be easily observed that, It's showing confidence level by fully analysing the image.

**Conclusion**: We have seen the most choozen dogs names, the most favourites a dog got on twitter and the confidence level of the neural network according to dog images.