

## Report: act\_report

What I found in this project is that people tend to give their dogs very high rating. They may have strong feelings for their dogs and that cause the users tend to rate a high score to their dogs. There are also some anomalies in this score, such as a score of 14 which is the maximum score in the dataset. Maybe the users should be told what the range of score that they should input, so that the data collection will be more accurate and reduce the amounts of outlier. There were also data problems when analyzing the most common dog names, such as 55 dogs with the name 'a'. It is possible that the user entered the wrong data, input 'a' as their dogs' names, but regardless of the strange data, we can see that Lucy, Charlie, Oliver, and Cooper are the most popular names.

Moreover, using neural networks to run out the picture analysis is a very big feat, and the algorithm can help us to massively identify every picture of a dog. From the first prediction, we know that the algorithm was able to identify the content of the photos as dogs with 73.8 percent. That's a stunning result for a developing algorithm. In another hand, the algorithm can be optimized to increase the recognition accuracy. Not only to identify the dog as the content, but also to identify the types of dogs, which could be very challenging task for programming developers.

Lastly, when I was visualizing the data, I used color-coded bar charts to separate the different dog names so that the readers could better identify the different dog names and the names' frequencies. The bar chart provides a better comparison of the frequency of different names. In the chart, we can clearly see that the most common dog names appear at very similar frequencies, with a frequency of around 10 times.

