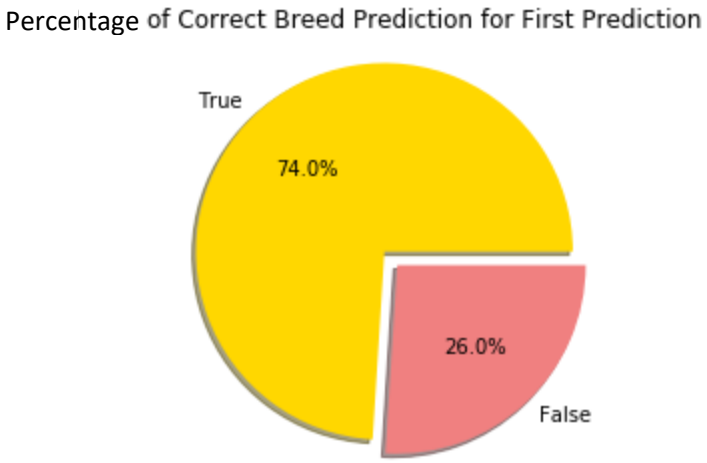
Who doesn't love a cute dogs with cute names? This project was done to analyze how an algorithm did in predicting the breed of dogs for each prediction and also to know which name of dog got more favorites count.

After gathering data from 3 sources, I assessed this data to know what I could find and detect cleaning issues. Yes, I cleaned it and got down to analyzing the data to satisfy my curiosity.

Analyzing whether or not the 1st prediction is a breed of dog was my first point of action. With values for this column being 'True' or 'False', I found out that approximately 74% True and 26% False. For the 2nd prediction there were 75% True values and 25 False values. In the 3rd prediction, The False values increased to 28% with True values at 72%.



Names of dogs might be an issues for some owners (for people like me), I was curious to know the name with more favorites count and created a list of the first 10 based on favorites. The dog with the most favorites count did not have a name to my disappointment but there were lovely names like; Kenneth, Pipsy, Hurley, Rubio and Terrance. 50% of the dogs in my top 10 list had no names (which might have been due to incorrect inputs which I had to clean during the cleaning process) leaving 50% with names. One of the constraints I faced with gathering this data for retweets count and favorites counts was that, they didn't match all the tweet IDs I needed because these tweets were deleted or not available.

I would be working on this data later on and would find more insights but for now, I might be thinking of calling my next dog Hurley. What about you?