Nicole Aguilera
DSC 680 - Milestone 2 Check-in

1. Not much has really surprised me from this data set. Something that has surprised me about the general project, though, is how accurate it was able to get. Looking at dogs, I as a human can tell the popular breeds apart from one another. But getting a computer to do this seemed pretty tricky. There are times when a dog can be a totally different breed than what they look like. So far with my project, the model has been able to correctly identify the dog breeds that the images are classified as, so I am very pleased with the results.

2. The data set is full of images. The one thing that did surprise me about this data set is that there are only 14 breeds included. I think that if I were to create this data set, I would include so many more breeds, maybe some of my personal favorites. The American Kennel Club has established over 100 breeds, so 14 is barely scratching the surface of these types of images. It would be very easy to include more images into this data set. I think that the dogs in the new images, however, would have to be of pure breed to have the model work correctly. If mutts were also included, the prediction would most likely not be accurate because it would only identify the breed that the dog looks the most like.

3. I have not adjusted my research questions or approaches to the model in question. It seems to me that the question is simple and straight forward enough that can be applied to an image classification model. I guess one question to go along with the project, but is not necessarily a research question, is whether or not the users should be allowed to put in any image that they want to. If this were the case, the model might return a dog breed for an image of a frog. This concept is further assessed in the following questions.

4. My method for this image classification problem is working. The model was able to fit to the data with a result of almost 97% accuracy. I think that this is a good ending point for the level of accuracy that it achieves. For the next milestone, I would like to test it on an image that is not a dog and see how it is classified with only the 14 breeds that it is trained on. I expect that the model will either return an error, and say that the image cannot be classified, or it will return a random breed because it cannot identify anything.

5. The challenge with this project is that if this model were to go into production, how would the user know to input images of dogs that resemble or that are of the same 14 breeds? The user can put an image of a dog that is not within those 14 breeds, but it most likely would not be very accurate. Again, it would be easy to put in more breeds, but I do not have the knowledge of putting together these images. If I had the time and resources to do so, I would add at least 16 more to have a total of 30 breeds to be run in the model.