

Module 2 Lab Exercise - Journal - Elijah Ghaya

To start this exercise I first downloaded the given files from canvas and imported it into google colab. I then carefully went through each cell and did my best to not only run the cells but to understand what exactly was going on in each line of code.

In part 1 the main objective was to import all the libraries such as Pandas, Numpy, and Matplotlib. Then the goal was to print the versions of Pandas and NumPy.

In part 2 the objective was the load in the datasets that we are going to be plotting and charting. This part was easy since the dataset already exists and the only thing needed was to bring it in. The print statements in this section were to display the shape, features, and the classes of which will be plotted and charted. This section was where I was able to get my answers for part of the reflection questions. Such as the amount of samples, features, and classes throughout the dataset.

In part 3, the focus shifts to visualizing the data using Matplotlib. Using Matplotlib I created a scatter plot to see how the lengths and widths of the different classes across different Iris species. This part was a main learning point for me because the data finally turns into a visualized picture. Here is where I could also see the most differences between the species and where I learned that the Setosa class is very different from the other two more similar species.

Part 4 was about practicing with basic data operations. This part was more about counting the the values of each species and creating a mean for each statistic. Here we learn that there are 50 samples for each species equaling out to 150 samples in total. In the assessment section I also added a line to display the total number of samples in the dataset, which might not mean a lot but definitely helps keep the data execution organized.

Overall this was an easy but very informative assignment. I learned new things and I feel a lot more comfortable with using datasets and creating visualizations for the data. I am looking forward to applying these same techniques to more complex datasets in the future to see what other patterns I can discover.