PDAT 617 Assignment 6 Due on July 20 11:59pm

We are going to implement k-means clustering using a famous dataset: the Iris dataset. This dataset contains 3 classes of 50 instances each and each class refers to a type of iris plant. The dataset has four features: sepal length, sepal width, petal length, and petal width. The fifth column is for species, which holds the value for these types of plants. (100 points)

Steps for training a model with K-means algorithm.

- **Step 1**: Load Data; Select all four features (sepal length, sepal width, petal length, and petal width) of the dataset in a variable called x so that we can train our model with these features.(30 points)
- **Step 2**: Using Elbow method to determine how many clusters you will choose.(30 points)
- **Step 3**: Build your model and import your k-means cluster center .(40 points)