Machine Learning   Project 1

Load Iris data set from the UCI machine learning repository(https://archive.ics.uci.edu/ml/datasets/Iris). Save the file as "Iris.csv" on your computer. Open your file in Excel, insert a new row at the top, type header names (feature names) and save it (make sure it is a ".csv" file). Then, write a python program to read Iris data set from the ".csv" file and answer the questions below:

a) How many tuples does the Iris data set have? (Check out the data frame)
b) How many features does the Iris data set have? Print a name list.
c) What's the data type for each feature?
d) How many of each species do you have?
e) What's the mean, median, quantiles and range (max-min) for each petal and sepal measurements?
f) Make a scatter plot of `sepal_length` vs `sepal_width`.
g) Create four individual histogram plots in a single figure, where each plot contains one feature (`petal_width`, `petal_length`, `sepal_width`, `sepal_length`). Hint: use Pandas library
h) Create four individual boxplots in a single figure, where each plot contains one feature (`petal_width`, `petal_length`, `sepal_width`, `sepal_length`).
i) Make a scatter plot matrix to examine the correlation between each of the measurements.

Note:

Your python program should be **commented**, and **structured**. Please place your file(.ipynb) in a directory named after you, zip it and submit it to canvas.

Projects that are demonstrated successfully live during class(at the end of class, last 15 mins of lecture) will receive an immediate of 100%.