Instructions:

Please read the instructions very carefully.

1. This is an **INDIVIDUAL WORK** unless specified otherwise. Students are not allowed to share their answers (actual coding) but are allowed to discuss with each other to solve the problems.

2. The tasks will be checked during the practical session itself therefore students will be assessed based on participation and their answers.

3. To start:

    a. Create a new folder called **StudentID-P1**. Please change student id to your own personal ID and change the 1 to the correct practical number.

    b. Create one .ipynb / .py file for each question with the naming convention question1.ipynb, question2.ipynb and so on.

    c. The following information must be included in each file:
        i. Student Name
        ii. Student ID
        iii. Module Code and Title

4. Please note that **ALL FILE AND NAMING CONVENTIONS** must be followed.

5. The **GREEN** colored font in the sample output represents an input from the user.

6. The **BLUE** colored font in the sample output represents a dynamic output.

7. Please note that the colored fonts will vary on the values specified. In other words, they are just **SAMPLES** only.

8. All tasks must be completed within the session. Students are given enough time to complete the tasks listed.

9. Once completed, please create a zip file with the same name as your folder, and upload it to LMS before the end of the session. A submission link will be provided.

10. Students are encouraged to upload their work to their own GitHub account. Students are also encouraged to include the GitHub link in the submission.

11. Students are encouraged to ask questions during the practical if they encountered a problem.

Question 1:

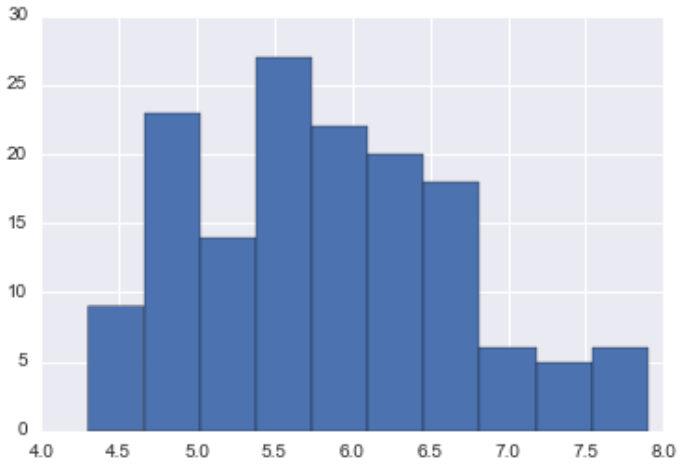
a) Apply data visualisation techniques on iris datasets. For each visualisation, you are required to describe each plot. The visualisation required are:
   i. Histogram of each attribute
   ii. Pair plot
   iii. Scatter plot (sepal-length against petal length)

b) Perform data visualization techniques on any datasets of your choosing. Use Question 1 (a) as a guide. You are encouraged to use additional visualisation aside from Question 1 (a) to help visualise your dataset. For each visualisation, you are required to describe each plot.

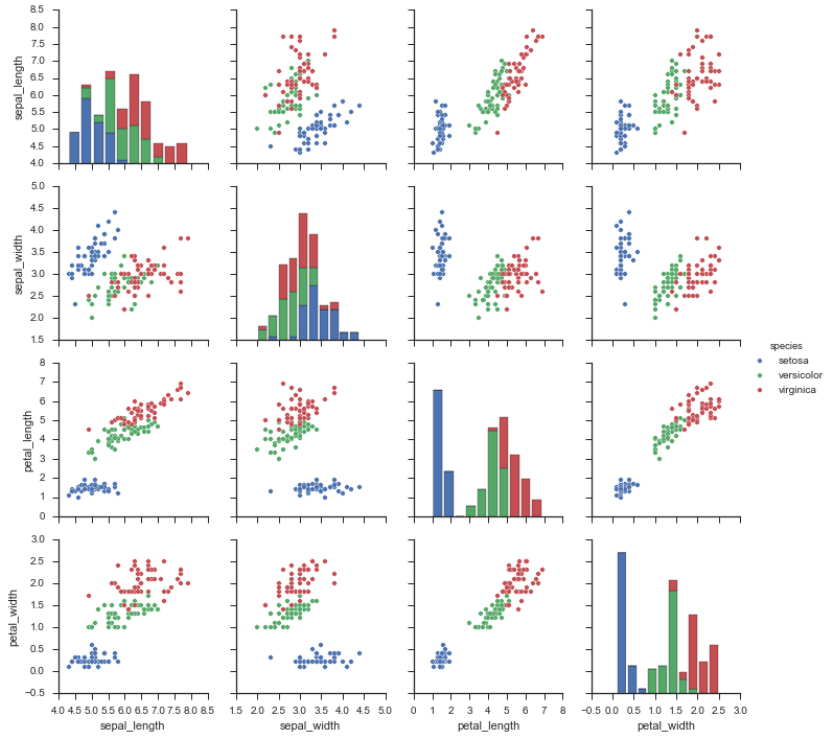
Sample Output:

Note:
i. The colour scheme of the visualisation can be change according to your preferences.
ii. Descriptions should also be included in the notebook

i. Example description of sepal length (other attributes you should complete it on your own): The highest value is around 8cm and the lowest is at 4.5 cm for sepal length. Most of the datasets falls in the range of 5.5 cm to 6.5cm.

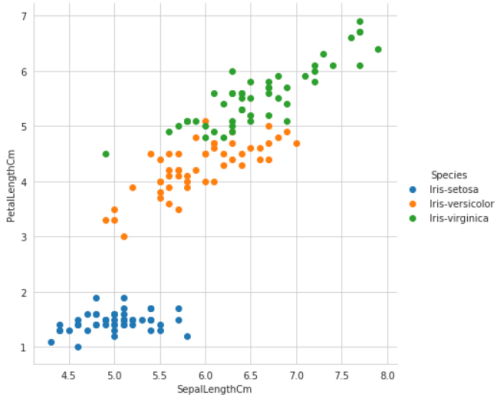


ii. Example Description: This shows all attributes of the datasets. This is to check if the iris dataset can be classify using only 2 attributes/features. It shows that a pair of sepal length and petal length, a pair of sepal length and petal width, a pair of sepal width and petal width and, petal width and petal length can partially classify the iris dataset.



iii. Example description: The plot shows the petal length and sepal length of the iris. From the visualisation:

- it shows that if the sepal length is greater than 7cm and the petal length is greater than 5 cm, it is considered as Iris-Virginica.
- It also shows that if the sepal length is lesser than 5 cm and the petal length is lesser than 2 cm, it is considered as Iris-Setosa.
- It also shows that if the sepal length is ranges from 6 to 7 cm and the petal length is between 4 and 5 cm, it is considered as Iris-Setosa.
- Other than the stated cases, it is diffcult to separate between the three types of iris.



- END OF FILE -