Alexandria University
Faculty of Computer and Data Science
General Department

Data Visualization Assignment

Objectives

This assignment aims to make students familiar with MATLAB programming and to begin the exploration process of a dataset.

1) .Load Data

In your MATLAB script load the iris dataset.

Create a data frame for the iris data using the “table” built-in function in MATLAB.

	1 sepal_lengt	2 sepal_widt	3 petal_lengt	4 petal_widt	5 class
1	5.1000	3.5000	1.4000	0.2000	seto...
2	4.9000	3	1.4000	0.2000	seto...
3	4.7000	3.2000	1.3000	0.2000	seto...
4	4.6000	3.1000	1.5000	0.2000	seto...
5	5	3.6000	1.4000	0.2000	seto...
6	5.4000	3.9000	1.7000	0.4000	seto...
7	4.6000	3.4000	1.4000	0.3000	seto...
8	5	3.4000	1.5000	0.2000	seto...
9	4.4000	2.9000	1.4000	0.2000	seto...

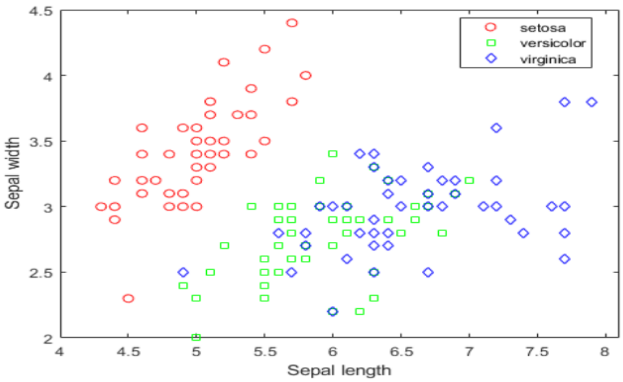
Alexandria University
Faculty of Computer and Data Science
General Department

Data Science Tools

Assignment

2) Visualization:

1. Plot the histogram for target data.
2. Use a scatter plot to plot every 2 attributes together.



3. Use 3D scatter plot to plot every 3 attributes together.
4. Create a new figure and define two “uipanel” objects to divide the figure into two parts. In the upper half of the figure, plot the sample data using “scatterhist”. Include marginal kernel density plots grouped by species. In the lower half of the figure, plot a histogram of the sepal length measurements contained in the target.
5. Show the correlation matrix between the attributes and comment on it.
6. Use boxplots to show the quartiles and the median of the sepal length, sepal width, petal length, and petal width for each kind of flower.
7. Repeat the project using Python programming.

3) Deliverables:

Submit a PDF report that includes the following:

- Explanation of the iris dataset attributes and details in a table form.
- The visualization and comments on them if available.
- Compare between Python and MATLAB programming and what you prefer and why?

4) Submission Instructions

1. Submit separate codes for the task and put them in a folder called “src”
2. Comment your codes sufficiently.
3. Keep the datasets in a separate folder called “data”.
4. Submit a README file that will contain the instructions on how to execute your codes and all source codes, report, result, and the README file compressed file (.tar.gz or .zip).
5. The compressed file should be named as [ProjectNo_IDs_Team_NUMBER.zip].
6. Submit it in the same email.

Palagrizim will be severely punished!!!!