COMP 1012 Summer 2018 Assignment 5

Due Date: Sunday, August 12, 2018, 11:59 PM

Material Covered

- Plotting
- matplotlib

Notes:

- Name your script file as follows: <LastName><FirstName>A2Q1.py. For example, LiJaneA2Q1.py is a valid name for a student named Jane Li. If you wish to add a version number to your file, you may add it to the end of the file name. For example, SmithRobA2Q1V2.py is a valid name for Rob Smith.

  Spyder automatically adds the .py part (called the extension) to your file name, so you don't have to type that part when saving the file. However, please ensure that your file name DOES include the .py extension, since our marking software looks for it.
- Name your output file as follows: <LastName><FirstName>A2Q1output.txt. For example, LiJaneA2Q1output.txt is a valid name for a student named Jane Li.

  In order to prevent Spyder from automatically adding the .py extension to your output file name, you must type the complete file name, including the .txt extension, when saving your output file.
- Follow the posted programming standards to avoid losing marks. Check your script for adherence to the programming standards by using CheckStandardsV2.py, which will show you exactly where your script does not comply. We will use the same script to assign marks for your assignment, so you have NO EXCUSE for losing these marks.
- You must complete the ***Blanket Honesty Declaration*** checklist in order to submit your assignment. This applies to all assignments in COMP 1012.
- To submit the assignment follow the instructions on the course website carefully. You will upload both script and output files, via the course website. We will demonstrate the assignment hand-in procedure in lectures. There will be a period of about a week before the due date when you can submit your assignment. ***Do not be late!*** If you try to submit your assignment after the late submission deadline, you will get a message indicating that the deadline has passed.

The time to submit this assignment is strict. Under no circumstances, the time can be increased. You will get 0 if you are late.

Question 1 [100 marks]

In this assignment, you are provided with two files: reducedDim.csv and conds.csv. The reduecedDim.csv file contains two columns and 16377 rows where each row indicates a point in two-dimensional space. The conds.csv file contains the labels for each data point in reducedDim.csv file. Your task is to create a scatter plot giving each label a different color and shapes.

Your program needs to have the following functions:

`def readData(fileName):` This function takes reducedDim.csv as parameter value and return a list named data. The length of data is 2 where each element in the data is also a list. The first element of data is a list containing all x-axis value and the second element of the data is a list containing all y-axis value.

`def readLabels(fileName):` This function takes conds.csv as parameter value and returns the labels as a list.

`def plotByGroup(data, conds, fmtstrings):` data is an array of shape (16377, 2) where the first column indicates x-axis and second column indicates y-axis. conds is an array containing all the labels. fmtstrings is a list containing 8 different colors and shapes that can be used to give each unique label a different color and shape. This function will plot the data with different color and shape based on the labels.

`def main():` This is the main function. From this function, you need to call the other three functions with appropriate parameters and values.

`Sample output:`

Hand-in

You will hand in your program script file. Also hand in the output produced by your code that looks like the example. You also need to submit the figure that you created. You can select your output in the IPython console window, then right-click and Copy (Raw Text) and paste it into a new document.