QQI

MSc AI

SUMMER 2022 EXAMINATION

Module Code: **B9AI108**

Module Description: **Programming for Data Analysis**

Examiner: **Paul Laird**

Internal Moderator: **Clive Gargan**

Date: Tuesday, 3 May 2022
Time: 09:30 – 12:30

INSTRUCTIONS TO CANDIDATES

Time allowed is 3 hours
Answer All Questions
Run each question as python in Google Colab. You MUST submit the .ipynb source file. Failure to submit the source will result in a zero grade. Please also screen shot the program and its output and paste in the Answer Sheet.

Question 1

a) Write a class in python, Maze, that represents a partially connected square grid, with attribute length (int) set to the value of the constructor argument n, and __data, an n*n array of integers, initialised in the constructor to numbers between 1 and 15, accessed by getter taking the index of the desired number.

The number may be interpreted as 8*[can move up] + 4*[can move right] + 2*[can move down] + [can move left]

A method, verifyPath should be included, taking starting coordinates and a list of integers in {1,2,4,8}, which should return the number of valid moves in the path (if all moves are valid it will return the length of the list), where 8=up, 4=right, 2=down and 1=left

(**15 marks)**

b) Write a shortest path solver function, which takes an instance m of Maze and determines a path from position {x,y} to goal {a,b}, where m,x,y,a,b are parameters to the function

(15 Marks)

c) Ensure solver returns the shortest path

(10 Marks)

(Total: 40 Marks)

Question 2

Data are provided on moodle in the file data.dat, which is in tab-separated format. Provide code snippets for the following:

a) Read the data into a dataframe and output a dictionary containing the sum of numerical columns, and the mode categorical columns, by column name.
b) Replace the values of feature B with a base2 log-transformation of feature B.
c) Present a scatterplot with feature C on the y-axis against feature D on the x-axis for the records where the value of feature F is the mode of feature F
d) Create one hot encoded binary features for each unique value of feature F.
e) Exclude the data where F is the mode, and create a test set of 20% of the remainder, randomly selected.

(5 * 5 = 25 Marks)

Question 3

a) Write a program in python to implement a programmer registration system using OOP concepts where programmers are assigned to teams. Create a separate team and programmer class. Each programmer should have an ID, salary and team; and each team should have a named manager. Your program must be menu driven. The user should identify themselves at the main menu and return to the main menu when finished.

Please screenshot sample interaction with your application

(35 Marks)

END OF EXAMINATION