University of Essex

School of Computer Science and Electronic Engineering

CE156: An approachable Introduction to Programming

Module Supervisor: Dr Giorgos Papanastasiou

Due Date: 26 March 2021 by 23:59 (UK time) via FASER

Assignment maximum mark: 100

Proportion to final module mark: 50%

SUBMISSION REQUIREMENTS: Students are required to submit ONE .py file PER QUESTION (i.e. as the assignment involves 4 main Questions, you have to submit 4 separate .py files) containing the code, for this assignment until 23:59 (UK time) on the 26/03/2021.

The standard lateness penalty will be applied to late work (contact the School Office for further information if needed).

Do NOT include any other project file (i.e. do not submit files other than the .py files).

FEEDBACK FROM THIS ASSIGNMENT Individual feedback will be provided via FASER approximately four weeks after the deadline.

Marking Scheme

Characteristics of an excellent project (70% or more):

• Excellent code documentation (i.e. descriptions using the “#” in your Python scripts)

• Excellent use of Python’s native methods, code standards, standard data structures and NumPy and/or Pandas dataframes where needed

• Excellent use of relevant data types

• Follows carefully the specification provided (where applicable)

• Excellent code optimisation in terms of result/outcome production and readability

• Generally, an excellent solution, carefully worked out, producing and demonstrating clearly and correctly results/outcomes

Characteristics of a good project (60%):

• Good code documentation (i.e. descriptions using the “#” in your Python scripts)

• Good use of Python’s code standards, standard data structures and NumPy and/or Pandas dataframes where needed

• Good use of relevant data types

• Follows the specification provided (where applicable)

• Good code optimisation in terms of result/outcome production and readability

• Generally, a good solution which produces and demonstrates correct results/outcomes

Characteristics of a fair project (50% or less):

• No meaningful code documentation

• Code tends to be lengthier and more verbose than needed, or at times difficult to read

• No real thought on the relevance of data types

• Does not follow the specification provided (results/outcomes are produced without addressing the question).

• A solution that only seems to deliver part of the results without showing some logical code developments.

Please note:

- You must submit **only one file per Question** (e.g. Question 1). Note: NOT one file per sub-Question (e.g. Question 1A).

- You are free to devise your method but you must follow the instructions (e.g. develop a **function** if the questions asks you to do so).

- You should **document** your code.

- It is a good idea to test each part of the code that aims to address a specific sub-question, separately.

Question 1:

A. Write a **function** that asks the user her/his age, makes sure it is actually a number and if that’s the case prints it on the screen. If it is not a number, just print “not a number”. Make sure that you write a function to perform the above. (12.5 marks)

B. Write a **function** which receives a list containing numbers as parameter. Your function should return True if all the elements in the list are in order, and False otherwise. Make sure that you write a function to perform the above. (12.5 marks)

For example, it would return True if you your list is:

alist=[0,1, 2, 3, 4, 5, 6, 7, 8, 9, 10]

and it would return False, if your list is:

alist=[0,1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 9]

Question 2:

Say that I have a dictionary “d” that I need to post-process:

d={'a':'1', 'b':'2', 'c':'3', 'd':'4', 'e':'1', 'f':'1’, ‘g’:’2’, ‘h’:’3’, ‘i’:’1’}

The above dictionary maps letters to numbers (frequencies). My main task is to develop a function that inverts the dictionary, i.e. maps numbers to letters.

Use a for loop to perform this. (25 marks)

Question 3:

Write a function that receives a string as parameter and then returns a dictionary mapping frequency to letters.

Eg. If the string is ‘hello’, the dictionary should be:

{1:['h', 'e', 'o'], 2:['l'],} (25 marks)

Question 4:

The NHS has given you a project to develop a data analysis task, which you will deliver in the form of a software (i.e. .py file).

The data of this project are housekept under an excel file (.xlsx).

A. Your first objective is to compute the total length of treatment for each patient. You can use Pandas dataframes and/or NumPy arrays to perform this. Print the results in a separate excel file called “Question4A”. (13 marks)

B. List separately the length of treatment for patients who use drug 390 vs. the generic. Print the results in a separate excel file called “Question4B”. (6 marks)

C. Compare the length of treatment by drug 390 vs. the generic by HR/HER2 status patients. Print the results in a separate excel file called “Question4C”. In total, there can be 4 different groups based on HR/HER2 status: (6 marks)

HER2(+)_HR(+),

HER2(+)_HR(-)

HER2(-)_HR(+)

HER2(-)_HR(-)