PROJECT 1 (15%)
STQD6014 DATA SCIENCE

PART 1 – INTRODUCTION TO DATA SCIENCES AND ALGORITHMS

1. What do you understand about data science?
2. Assuming you are a department manager, and would like to investigate the customer's preference on three types of your company's products. Hence, give a bit introduction about your company and what types of products you want to investigate. Next, explain the active roles of data scientist for this task.
3. Based on notes week two of "Basic of Algorithms", find/create one problem. You may refer to example of "Direction of numbered NYC streets algorithms" or "Class average algorithms".
   a) State the problem, input, processing and output
   b) From the problem specified in part a) above, create three popular program design tool (flowcharts, pseudocode and hierarchy charts). For flowchart, please include as many symbols as possible.

   Note: your solution must be unique.
4. Give two example of current data technology and its explanation.

PART 2 – LISTS, DICTIONARIES AND LOOPING

You are developing a product price management system for a store. The program should:

1. Continuously prompt the user to either add a new product, update a product's price, view all products, or stop the program.
2. Store product names as keys in a dictionary, with the values being a list containing the product's price and quantity.
3. Allow the user to add new products with their price and quantity.
4. Allow the user to update a product's price or quantity.
5. Allow the user to view all products with their prices and quantities.
6. Exit the program when the user types "Stop".
7. If an invalid input is provided, prompt the user to try again.

Write one page of essay (font 12 and spacing 1.5) explaining your inventory management system.

PART 3 – FUNCTIONS AND CLASSES

For this task, you have to use ***class*** and save it as a ***module***. Create a scenario involving function, class and inheritance.

1. Write the introduction of your scenario.
2. Write a program using python class, function and inheritance for the chosen scenario above.
3. Test your program by creating five instances of each parent and child class and print the info related.
4. Write the explanation of the program created above.
5. Provide your conclusion.

Write the item number 1, 4 and 5 in essay form. Your essay should be at least two pages long using times new roman, font 12 and spacing 1.5 excluding the Python codes.

PART 4 – FILES AND EXCEPTIONS

1. Make two files, movie.txt and sport.txt. Store at least five names of movies and sports line by line in these corresponding text files.
   a) Write a program that tries to read these files and print the contents of the file to the screen.
   b) Add five newlines to both of these files, describing the existing contents with suitable explanation.
   c) Wrap your code in a try-except block to catch the FileNotFound error, and print a friendly message if a file is missing. Move one of the files to a different location on your system, and make sure the code in the except block executes properly.
2. Write a program that prompt for user's year of birth and favourite number. Use json.dump() to store this number in a file. Write a separate program that reads in this value and prints the message,
   "My year of birth: "
   "My favourite number: "

PART 5 – NUMPY

Create two arrays each of sizes: single 1d array, 3x3, 2x4, 4x2 and 5x5. Hence, you will have ten arrays all together and these arrays must be generated via random number. Name your arrays in order from A to J.
Using numpy library: perform at least twenty (20) functions from Table 4-3, Table 4-4, Table 4-5 and Table 4-7 from Python for Data Analysis book.
You must:

1. Perform operation on single array
2. Perform operation on two arrays of same size (eg. multiplication, etc.)
3. Perform operation on two arrays of different size (eg. dot product between array of size 2x4 and 4x2

Note: you must include all arrays created.

Instruction:

1. For Task 2 to Task 5, provide different python notebook for each task.
2. Please submit your project in UKMFolio by 5th January 2025 before 5.00pm.
3. Penalty will be imposed for late submission.
4. Higher marks will be given for unique submission with minimal use of AI tools.