Deadline: 11:59 PM, Monday, 19th June, 2023
Submission: Push your code to github and email us the link to the repo in reply to this email

Instructions:

Complete the following assignments by midnight on the 19th of June, 2023. Once complete, push your code to Github, and email us the link to the repo. Include a readme file explaining your process.

Please note in your readme:

- The version of python and the environment setup you are using.
- If we need to install any libraries or dependencies in order for your code to work.
- Any setup that we must do for your code to run smoothly.
- Detail any additional features you have chosen to add for both the tasks.

Even if you cannot finish, please submit as far as you manage to finish, explaining what you struggled with and how you tried to solve it.

Assignment Part 1: Excel

1. Please download the attached excel file.
2. There are 2 sheets in this file.
3. Create a pivot table using data of sheet1 and show the information following this structure -
    a. The data table should show the Sum of Income as value.
    b. The columns should include the value of Gender and MaritalStatus.
    c. The rows should be in the following order: Division; Customer Name; ID.
4. In sheet2 there are some IDs. Add a new column to sheet1 and name it “Matched”. Please match the IDs of sheet2 with the IDs of sheet1 and show the result as True or False. You need to use a formula to match the IDs.

Note: Feel free to add additional sheets and show the data of sheet1 filtered by some parameters any value of the columns as a bonus point.

Assignment Part 2: Python

1. In this part, separate the sheet 1 data in a new excel sheet.
2. Convert the excel file to csv.
3. Now load the csv file.
4. Get rid of the column ID from the data frame.
5. Encode the data to have similar values.
6. Now use K-means clustering based on their divisions. This part is a bonus task.
7. Download the file without omitting the output.

Note: Use colab to do this task. Add the csv file and the ipynb file in the repo.