# Assignment

**Module**

Data Processing

**Module Code**

DE4003FP

**Duration**

5 hours

**Title:**

Apply data transformation and merging methods to a given dataset

**Tools, Equipment and Materials:**

- Personal Computer with Internet access
- Data Processing Software

**Instructions & Guidelines:**

1. Submissions should be made via the myConnexion by the stated deadline.
2. No marks will be awarded, if the work is copied or you have allowed others to copy your work.

   **Submission:**

3. Submission should include the following items:

   - Microsoft Power BI file
   - Microsoft Excel file

# Assignment

**Module**
Data Processing

**Module Code**
DE4003FP

**Duration**
5 hours

## Task
## Requirements

**Dataset:**
A research study has been conducted on different races on the common risk factor for heart diseases and type 2 diabetes. The dataset contain the history of hypertension, cholesterol level and BMI.

**In Microsoft Excel**

**Part 1:**

### Assign data types

1. Assign correct data type for these columns: Sex, Marital, Race.

### Perform data manipulation columns into tables

2. Use Martial column to extract two dataset as follows:

#### First dataset :

a) Show Martial column **without** blanks.
b) After that, transfer all the columns including Martial column without blanks into another spreadsheet, name it as first dataset.

#### Second dataset:

a) Show Martial column **with** blanks.
b) After that, transfer all the columns including Martial column with blanks into another spreadsheet, name it as second dataset.

3. Show Income and Sex from the highest to the lowest in the first dataset.

4. Insert a text box in the first dataset and write your answers for these questions.
a) What is the highest income?
b) What is the lowest income excluding the blanks?

### Replace row values in the dataset.

5. In the first dataset, replace the row values in the all columns as 0.

# Assignment



6. Use this cholesterol formula:

HDL + LDL + 20% of Triglycerides

Apply this calculation in a new column for the first dataset. Save the Microsoft Excel file.

**Part 2:**

**In Microsoft Power BI**

**Perform data profiling activity**

7. Load the Microsoft Excel file: first dataset into the Microsoft Power BI.

8. Display the data profiling for the first dataset and save the Microsoft Power BI file.

**Merge data into required tables**

9. Import the BMI file into Microsoft Power BI.

10. Clean the BMI data as required.

11. Display the table relationship between the first dataset and the BMI file. Save the Microsoft Power BI file.

**Create visualisation for the dataset**

12. Create a chart to show Age and average BMI in Page 1 in the Microsoft Power BI.

13. Create a chart to show the count of Patients and Race in Page 1 in the Microsoft Power BI.

14. Adjust the two charts in Page 1. Save the Microsoft Power BI file.