

In this assignment you are required to:

- Prepare and conduct stepwise Data Management activity such as Data Exploration, Metadata understanding, and Exploratory Data Analysis based on the data given. You are required to use any data analytics platforms (excluding Ms Excel Spreadsheet) to complete this assignment. Use the following data for your assignment:
 - i) SAS Viya 3.5 : You are required to use **HEART.data** (path: TUNDATA)
 - ii) Any Other platforms used : You are required to use data provided in class – **Player_FIFA22.xls**.
- Document the solutions as set out in the assignment requirements in a clear manner. The word limit (**excluding charts/tables/diagrams and references**) for the report is **800-1000** words.
- Submit your assignment documentation online via Moodle.

Exploratory Data Analysis and Data Pre-processing**Requirements:**

Examine the given dataset and carry out the following tasks.

1. Identify the type of each attribute (nominal, ordinal, interval or ratio).

[10 marks]

2. Conduct an **Exploratory Data Analysis** for the given dataset using SAS Studio / SAS Viya. You may include the following in your analysis.

- i. Descriptive statistics (e.g. value ranges of the attributes, frequency of values, distributions, medians, means, variances, and percentiles).
- ii. Suitable visualizations for the corresponding statistics.

[30 marks]

3. Use SAS and apply **pre-processing techniques** to handle:

- i. Incomplete data
- ii. Noisy data
- iii. Inconsistent data

Provide detailed explanations on the applied techniques. Show the resulting dataset and justify your approach in each case.

[60 marks]