Technological University of the Shannon: Midlands Midwest

Faculty of Engineering & Informatics

Semester 2 (100% Continuous Assessment)

Summer Session

Postgraduate Diploma in Engineering in Engineering Management
Postgraduate Diploma in Engineering in Engineering Management (Jan Intake)

Year 1

Data Analytics

External Examiner: **Dr James Byrne**

Internal Examiner(s): **Ms Fiona M. Walshe**

Instructions to Candidates:

Read the instructions provided carefully and provide a submission for all parts of all questions. Starter code, CA2-starter-code.R, data, CA2-data.csv, and data variables descriptions, Data-Descriptions.xlsx, are provided for this assignment. Place both the starter code and the data file in the same directory in your computer, alter the starter code as necessary and then step through the code to run your analysis/analytics steps. You can open the Data-Descriptions.xlsx file to view descriptions of the data variables provided in the csv file.

This assignment is worth 60% of your total marks for this module.

Time allowed for the Assignment: 2 weeks
No. of pages (including cover sheet, attachments/drawings): 3

Assignment 2:

Background - An Iranian Construction Company wishes to employ data analytics to estimate construction costs based on physical and financial construction input variables. The model must perform within +/- 500,000 Iranian Rial to be of use.

(a) Data Understanding & Presence of Code

Comment on the data types and quantity. Following your code block insertions (code can be copied from previous examples and modified as necessary), correctly interpret and provide commentary on the exploratory data analysis that you conduct. As a final step, copy and paste your code from the project into this report's Appendix A (do this after you have answered the remaining sections).

(20 marks)

(b) Data Preparation

Following from the data understanding phase, you may have identified a data issue or potential improvement. Prepare the data to address this issue or implement the data improvement correctly. Make note of this issue or potential improvement here and provide a justification for what you have done.

(20 marks)

(c) Modelling & Evaluation

Use the appropriate data partition for model training and/or validation and do not bias your final test set result with data leakage. Accurately report on performance here with an appropriate measure or measures. An example table with example table headings is provided below in case you find this useful to highlight the results achieved. Provide a final evaluation of the model usefulness, based on the obtained results.

Table 1: Example Table Heading Here

Method	Performance	Notes

(20 marks)

[60 marks]

Appendix A

(Copy and paste your code below this heading. Do not worry about formatting of the code after it is pasted)