MA30280 Applied data science – Coursework 2 instructions

Set: 15 April 2024.
Due: 26 April 2024, 11.59pm.
Estimated time required: 20 hours for a student who is up-to-date with the unit materials.
Value: This coursework constitutes 50% of the unit mark.
Length: There is no specific page limit, but very lengthy solutions that fail to emphasise key points or get lost in unnecessary details will not receive a high mark (see Marking scheme below for more details).

Submission: On Moodle.

- Your submitted solution must be a compiling RMarkdown (.Rmd) file. If any additional resources (Excel files, images, etc.) are necessary for the solution, upload these to the submission point in a single zip file.
- The Rmd file should be named "MA30280_CW2_CandidateNumber.Rmd" and (to guarantee anonymity of the candidate) must NOT contain the author information in the YML header.
- The RMarkdown output (i.e. pdf, html, word) should include only the computer commands/output that form part of your solution to the questions.
- The R code chunks in the Rmd file should be commented. This will not be marked for credit, but it might be checked to see if it is unclear what you have done in your solution, so ensure that it is intelligible and does not contain redundant material.

Conditions: The work submitted is individual and personal. Your lecturer may answer generic questions about methodologies and computing relevant to the coursework, but not specific questions about specific analyses. To keep things fair, questions relevant to the whole class will not be answered individually, but will be answered on the Moodle forum. Do not ask other members of staff, post-graduates or students for help.

Support and advice: For any questions on the coursework assignment, contact me during the Q&A sessions in weeks 9 and 10, or via email at eahd20@bath.ac.uk. For IMCs or extensions, contact your Director of Studies.

Feedback: You will receive feedback for this coursework after the final mark for the unit is released in July. The feedback will consist of your marked work and an overall feedback document commenting on the assessment across the cohort.

Late submission of coursework:
If there are valid circumstances preventing you from meeting the deadline, your Director of Studies may grant you an extension to the specified submission date, if it is requested before the deadline. Forms to request an extension are available on SAMIS.

- If you submit a piece of work after the submission date, and no extension has been granted, the maximum mark possible will be the pass mark.
- If you submit work more than five working days after the submission date, you will normally receive a mark of 0 (zero), unless you have been granted an extension.

Academic integrity statement: Academic misconduct is defined by the University as "the use of unfair means in any examination or assessment procedure". This includes (but is not limited to) cheating, collusion, plagiarism, fabrication, or falsification. The University's Quality Assurance Code of Practice, QA53 Examination and Assessment Offences, sets out the consequences of committing an offence and the penalties that might be applied.

Generative AI: Type B
Generative AI is permitted as an assistive tool for specific defined processes within the assessment and its use is not mandatory in order to complete the assessment. In particular, under the University's Academic Integrity Statement, you "must not present content created by generative AI tools as though it were your own". Any text or code produced by generative AI must be checked for correctness and cited. You should be prepared to explain anything in your submission to an examiner if asked to do so. (See GenAI Assessment Categorisation).

Note: You must include a short statement (max 250 words) at the end of your submission indicating what tools you used and how you used them OR that you have not used generative AI tools.

Marking scheme:
This coursework will be marked out of a total of **50 marks** (a more detailed breakdown is given in the question sheet). There is often no single correct answer to the questions and the marks will be allocated on a combination of the approach taken, the justification for this approach, the interpretation of results in context, and the overall presentation.

35-50 (First) A solution that could be presented with little or no revision. Analysis is sound so that conclusions are scientifically well-supported. Interpretation is reasonably mature. The solution should demonstrate a clear overview of the work, without getting lost in details, and be free of all but minor statistical errors.

30-34 (2.1) A solution that could be presented after a round of revision, but without having to re-do much of the actual analysis. Some flaws in the analysis or presentation (or minor flaws in both), but basically sound. A good grasp of the modelling techniques and context, so that interpretation is reasonable.

25-29 (2.2) Major re-working required before the solution could be presented but containing some understanding of the modelling approach and its application. Reasonable presentation and organisation.

20-24 (Third) Major flaws in analysis and presentation, but demonstrating some understanding of methodologies, and a reasonable attempt to present the results.

0-19 (Fail) Flawed analysis demonstrating little or no understanding of methodologies, and/or incomprehensible or overly bad organisation/presentation.