|  |  |
| --- | --- |
| Task/assessment brief: | |
| **Overview**  In this assignment you are required to select a dataset from Kaggle, perform a rigorous analysis of the data and critically reflect on the chosen approach for data analysis. You are free to choose any dataset depending upon your interest but that needs to be approved by a member of the teaching team before you start working on the data. Additionally, to improve your analytical and problem-solving skills, you need to select an appropriate research problem and propose the idea to solve that problem. The subsequent sections will provide all the necessary details regarding the dataset approval process along with the major steps to be followed. This assignment is aimed at assessing your ability to pose interesting analytical questions, process the data using the key steps of data analytics, such as, data pre-processing, analysing, eventually preparing a reflective report, and problem-solving skills. In order for your analysis to be compelling it must address a substantive issue rather than a trivial one.  **Dataset approval**  Although you can choose any available dataset from Kaggle for this assignment but that must be approved by your workshop tutor. The approval should be obtained before the end of Week 6, which is the week starting 6 March 2023. If you have any problems in meeting this deadline, then please contact your module leader for discussing your situation.  **TASK 1 (Analytical question)**  Propose an interesting analytical question that can be answered using the given dataset. For example, let us consider a dataset, New York City Airbnb data available on Kaggle. An example analytical question could be – Do surrounding Airbnb listings impact a property's rental price in New York City?  **TASK 2 (Data analysis)**  In this task, use your programming-driven data analytics skills to answer the question posed in the Task 1. Depending upon your chosen question, you will typically have to perform Exploratory Data Analysis (EDA), data pre-processing, statistics/computation-based data analysis, and visualisation of the key results.  **TASK 3 (Critical reflection report)**  In this task, you are expected to critically analyse all the chosen methods employed in your data analysis. You must reflect on the rationale behind choosing those methods, challenges encountered during their implementation, method’s effectiveness in relevance to the analytical question, and any lessons learnt for the future works. The maximum word limit for this task is 1500. The word limit is applicable only to the actual review and the references/bibliography are excluded from this word limit.  **TASK 4 (Analytical and problem-solving skill)**  In this task, you are required to find the appropriate problem and their possible solution (idea to solve problem). The problem statement should be related to the content of this module. You are free to choose any application area. This assignment is aimed at assessing your analytic and problem-solving skills. For better presentation, you can show the problem and idea/solution using appropriate block diagram and text. The maximum word limit for this task is 500 words. This assessment will be evaluated based on novelty of problem and their appropriate solution (quality of idea).  **Assignment deliverable**  Your solutions to the Task 1 and Task 2 should be presented using a SINGLE Jupyter notebook (.ipynb) describing all the steps performed for achieving the objectives of both the tasks. All the pieces of codes in the notebook should be adequately commented and reproducible. The solution to the Task 3 should be presented through a SINGLE PDF document. The solutions to the Task 4 should be presented using a SINGLE word file. Please refer to the Submission Details section for more information on the assignment deliverable. | |
| Word count (or equivalent): | 4000 |
| This a reflection of the effort required for the assessment. Word counts will normally include any text, tables, calculations, figures, subtitles and citations. Reference lists and contents of appendices are excluded from the word count. Contents of appendices are not usually considered when determining your final assessment grade. | |

|  |
| --- |
| Academic or technical terms explained: |
| **EDA-** Exploratory Data Analysis  **Statistical Description-** Use of statistical parameters like mean, variance, standard deviations, etc. to analyse the result.  **Key Bloom elements:**  Comprehension – Demonstrate key aspects needed for the EDA  Analysis – Comparing and contrasting different Machine Learning Methods for a particular  Synthesis – Building the EDA around outcomes  Evaluation- Evaluate the results generated by different algorithms  Application- Apply methods in different application environments. |

# Submission Details

|  |  |  |  |
| --- | --- | --- | --- |
| Submission Deadline: | 12 May 2023 | Estimated Feedback  Return Date | 2 June 2023 |
| Submission  Time: | By 4.00pm on the deadline day. |  | |
| Moodle/Turnitin: | **Any assessments submitted after the deadline will not be marked and will be recorded as a non-attempt unless you have had an extension request agreed or have approved mitigating circumstances. See the School Moodle pages for more information on extensions and mitigating circumstances.** | | |
| File Format: | The assessment MUST be submitted using Three files through the Turnitin submission point on Moodle. The FIRST file should be a zip/compressed file within which all the separate files (.pdf/ .ipynb etc.) corresponding to the tasks (Task 1, Task 2, and Task 3) will be merged as a single zip file. The SECOND one should be a PDF file corresponding to the Task 3. The THIRD one should be a WORD file corresponding to the Task 4.  Your assessment should be titled with your Student ID Number, module code and assessment id, e.g., st12345678 CIS7031 WRIT1 | | |

# Assessment Criteria

|  |
| --- |
| Learning outcomes assessed |
| |  | | --- | | **Learning Outcomes** | | This assessment is designed to demonstrate a student’s completion of the following Learning Outcomes:   * [LO1] Critically analyse and evaluate various statistical and computational techniques for analysing datasets and determine the most appropriate technique for a business problem. * [LO2] Critically evaluate, develop, and implement solutions for processing datasets and solving complex problems in various environments using relevant programming paradigms. * [LO3] Evaluate and apply key steps and issues involved in data preparation, cleaning, exploring, creating, optimizing, and evaluating models. * [LO4] Evaluate and apply aspects of data science applications and their use. |  |  |  | | --- | --- | | **Assessment Criteria** | **100%** | | Task 1 – Posing an interesting analytical question relevant to the chosen dataset (LO1) | 5% | | Task 2.1 – Exploratory data analysis (LO1, LO2) | 15% | | Task 2.2 – Data pre-processing (such as, data cleaning, data transformation etc.) (LO3) | 10% | | Task 2.3 – Statistics/computation-based data analytics (LO3) | 30% | | Task 2.4 – Visualisations of the key results (LO3, LO4) | 10% | | Task 3 – Critical reflection report (LO4) | 20% | | Task 4 - Posing an interesting analytical and novel research problem with creative idea (LO1, LO4) | 10% | |
| Other skills/attributes developed  This includes elements of the Cardiff Met EDGE (Ethical, Digital, Global and Entrepreneurial skills) and other attributes developed in students through the completion of the module and assessment. These will also be highlighted in the module guidance, which should be read by all students completing the module. Assessments are not just a way of auditing student knowledge. They are a process which provides additional learning and development through the preparation for and completion of the assessment. |

|  |
| --- |
|  |

Marking/Assessment Criteria

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tasks | **<35%** | **35-39%** | **40-49%** | **50-59%** | **60-69%** | **70 – 100%** |
| Task1 | Obvious and trivial analytical question with no practical importance. Demonstrate no understanding of the given data. Significant error showing little time and effort in the preparation of the question. | Unfocused and unspecific analytical question. Contain some error but there is evidence to show that some effort has been spent in the question preparation | Substantive, but non-specific analytical question and/or overly ambitious. Reasonable evidence to show understanding of the selected topic. Contain some errors, but there is evidence to show good amount of effort have been spent in the question preparation. | Substantive and focused analytical question. Clear degree of understanding of the given data. Contains a few errors, appropriate for the amount of time and effort spent in the question preparation. | Substantive analytical question. Theoretically interesting with some practical importance. Demonstrates high degree of understanding of the given data. Contains very few minor errors demonstrates good amount of time and effort spent in the question preparation. | Specific and substantive analytical question which is theoretically interesting and have significant practical importance. Demonstrate an excellent understanding of the given data. Error free content, which shows that significant amount of time and effort was expended in the question preparation. |
| Task2 | Limited/no progress towards a solution to the specified question. Unprofessional solution. Does not demonstrate understanding of the problem and/or solution. | Progress toward a basic solution, but incomplete or inappropriate. Data analytics attempted, but incomplete or with significant errors. Inadequacy of in-depth skills, knowledge and understanding, with significant flaws in execution. | Adequate but basic solution but, with deficiencies evident. Basic design which addresses the key requirements. Demonstrates some skills, knowledge and understanding. Data analytics executed without full consistency or with flaws. Contains some errors and/or omits significant elements. | Good working solution presented to specified question, but with minor omissions evident. Demonstrates a basic level of competence and some in-depth skills, knowledge and understanding. Data analytics executed in a largely appropriate manner with some appreciation of best practice. | Complete solution to the problem with only very minor omissions. Clear and appropriate solution design. Data analytics executed and presented in a fully systematic and logical manner with an appreciation of best practice. Ability to apply knowledge gained during the course thoughtfully, accurately & effectively. | In-depth, and complete, solution to the chosen question. Complete in-depth design, following required standard with no errors. Data analytics executed and presented with independence in a fully systematic and logical manner, demonstrating clarity, professionalism, and best practice. |
| Task3 | Unclear structure to the report. Little/no consideration given to presentation of completed work. Assignment instructions not followed. Text has an unclear message. Little/no effort made to follow instructions. | Some attempt made at structuring subject specific reflections.  Some structure to report, but presentation / design requires more consideration. Many spellings mistake. Basic sentence construction rules followed. Some text may be verbose. Minimal ability in communicating ideas. Attempt made at following submission instructions, with major errors or significant omissions. | Critical reflection on task attempted, with minimal depth. Demonstrates minimal ability to learn through active reflection on work undertaken  Clear structure to the report. Some consideration given to presentation of solution. Some spelling mistakes in the report. Basic sentence construction rules followed. Adequate level of ability in communicating key ideas and concepts. Generally informative writing style. Submission instructions mostly followed, with errors / omissions. | Critical reflection on task attempted, with limited depths. Demonstrates ability to learn through active reflection on work undertaken. Well-written, clearly designed analytical report. Careful consideration given to presentation of solution. Some spelling mistakes. Most sentences of appropriate length and punctuation used. Generally informative writing style. Competent level of ability in communicating knowledge. | Good level of critical reflection.  Demonstrates  ability to  through  reflection on work undertaken.  Considered structure to report. Thoughtful, clear presentation of solution. Few / minor spelling mistakes. Clear, well punctuated sentences. Writing style clear and informative. Demonstration of high level of competence in communicating knowledge. | Excellent level of critical reflection. Clearly demonstrates ability to learn through active reflection on work undertaken.  Well-formulated structure to report. Professionally presented solution. No spelling mistakes. Clear, well punctuated sentences. Writing style professional, clear and informative without being verbose. Excellent skills in communicating knowledge. |

