# Assessment Background/Scenario

The demand for rental housing decreased from 12.3% to 11.1% in mid-2022, and is predicted to further decline to 4.5% by the end of 2024 [1]. A company housing manager is concerned by this downward trend, and you have been assigned tasks to identify and investigate three problem areas and develop potential solutions to these problems. You are required to utilise the data mining techniques (regression/classification) and tools (WEKA version 3.8.5) that have been taught in the Big Data Analytics module and only use the “Housing” data set provided, which can be cleaned and used to generate specific output.

### Data Set (.CSV)

The data set titled “Housing” is provided under License CC0: Public Domain. The data consists of two CSV files – housing\_train and housing\_test – one suitable for training and the other for testing.

**Reference**:

[1] R. Donnell (2023, Nov. 3). *Rental Market report: what’s happening to rents?* [Online]. Available: https://www.zoopla.co.uk/discover/property-news/rental-market-report-march-2023/ [Accessed: Nov. 3, 2023]

# Assessment Tasks

1. **Rental “demand” investigation: (ML02 & MLO3) (40%)**

The housing manager has the following question: which characteristics of a property determine the level of customer demand? To answer this, the manager proposes looking into the following:

1. Which of the “discrete variables” (e.g. bedrooms, smoking\_allowed) have the potential to predict a “low demand” property? Do these variables also have the potential to predict a “high demand” property?
2. Ascertain if there is a correlation (either positive or negative) between the “demand” for a property and its “rent” and “type”.
3. Identify if the size of the property “sqfeet” has an optimal range for generating high “demand”.

You should utilise Weka and build a classifier or regression model to perform this analysis.

1. **Storing data and scalable solutions: (MLO1 & MLO4) (40%)**

Part 1: Design a relational database

The housing manager is considering an alternative to the current flat file (CSV) system that stores the majority of their data. You have been tasked with designing a relational database to store the provided dataset ‘Housing’ in the flat file system. You will need to decide, and justify, which features to include and/or adapt to store all the provided data. To allow the housing manger to assess the feasibility of this you should provide the following:

1. Produce a database design (in the form of a UML standard ER diagram with normalisation to 3NF) for the given data.
2. Present sample SQL for the database you have created (given your ER diagram diagram) as follows:
3. Demonstrate the SQL that you would write to enter a new line of data, covering all relevant attributes.
4. Extract the ‘description’ for all properties with a rent equal to or less than 1000\*, allows both cats and dogs, and is in the state represented by ‘ca’.
5. Extract the average rental value for each state so they can be compared.

\* No currency is specified in the dataset.

Part 2: Consider scaling

The housing manger is also considering longer term solutions for their business, given their intention to set up international offices across the globe. This would generate considerably more data (tens of megabytes). To be able to utilise this data effectively requires a rapid-response system for the business to be responsive in a global rental environment. Assume certain messages are required to be sent as soon as certain automated analysis results are returned of a certain value (e.g. a count of items of a particular type exceeds a pre-determined threshold). With reference to specific details in the data set, present a way that you could use appropriate technologies to spread the load over multiple computers and justify why this would be a good approach.

1. **Considering public-facing application: (MLO5) (20%)**

The housing manager is considering the development of a public-facing application, to assist in promoting the expansion of his business, and to make it easier for potential clients to view and select from current offerings. As part of this, he is also considering capturing (via an online form) the personal details of potential clients so he can provide recommendations to revisiting clients.

Identify the three most salient privacy issues that he needs to consider before embarking on this new venture. You should consider the potential issues in the context of the new application, the rental agencies’ intentions regarding data analysis (Task 1), and the move towards permanent data storage (Task 2). For each of the issues you have identified, discuss the strategies that could be employed to address each of them in the context of the given scenario and data set.

# Deliverables

Your assignment should be laid out following the formatting guidelines that are specified in the **‘Submission Formatting’** page in Canvas. This includes restrictions on the length of the appendices, expectations on how your work should be presented and any penalties when these expectations are not met.

* Your submission should consist of a **report and final data files in ARFF** **format**. ARFF files should be submitted as part of a .zip archive.

Given the tasks above you should produce a report detailing solutions and justifying your decisions. You will need to provide supporting evidence for each solution in the form of images/screenshots of the practical work you should have undertaken to complete this assessment, including anything that is specifically requested. Your decisions and justifications should be supported by the current literature.

Your report should not exceed **3,000 words** in total and consist of three clear sections – one for each task. Your response to one section/task will **not** contribute to grades in another. Further formatting details and essential points are given below.

## Task 1: Rental “demand” investigation (MLO2 & MLO3 40%)

(Suggested word count for this section: 1500, i.e. 500 words per solution)

You can import the given data files into a spreadsheet to initially scrutinise and review the data, as well as perform any cleaning before you translate it into the .ARRF format for use in Weka. A notepad application can also be used to do this.

You are required to submit the .ARRF (or files) that you have used to perform the required analysis so we are able to verify your results where necessary.

Your report should discuss and present the following for this task:

* Any assumptions you made about the scenario or areas of investigation.
* Any pre-processing you have undertaken to make the data fit for purpose.
* Clearly state the specific analysis techniques you have employed in your solutions.
* Justification for the selection of techniques/approach, given the nature of the data and the requirements of the investigation, which is effectively supported by the literature.
* Provide a general summary of the results of your analysis, along with the specific results (in the appendix).
* Consider how your results, individually and as a whole, answer the question posed by the housing manager.
* Critically evaluate the approach you have taken, and the techniques selected in the context of the given data set and scenario. You should reflect on what you have learned from the process and identify what was effective/ineffective. An honest appraisal of an ineffective approach will gain credit. This discussion should be supported by the literature.
* For each solution you should provide images/screenshots that demonstrate the tool you have selected in Weka, any relevant settings, and the output produced by that tool.
* For each solution you should provide (an) additional file(s) containing the final data structure you used in Weka. This should be in Weka’s .ARFF format.

All **diagrams and images/screenshots should be presented in the** appendices which **must** be referred to and discussed in the body of the report.

To attain any **Grade** in this task you **MUST** present evidence of your work in Weka and the final data files you used to undertake this in .ARFF format only. This is required to verify your results, and therefore your discussion. Failure to do so will result in a grade of **zero** for this task/section.

## Task 2: Storing data and possible solutions (MLO1 & MLO4) (40%)

(Suggested word count for this section: 1000, i.e. 400 words for Part 1, 600 for Part 2)

Part 1:

You can either type out SQL statements OR build and screen capture (image) the SQL from a live database. You should present the ER diagram and the SQL statements in the appendices of your report and refer to each in your brief discussion of your approach.

You may choose to demonstrate the normalisation process using specific examples of your approach. Any further visual aids (tables/models/diagrams etc.) should also be presented in the appendices, not in the main body of the report.

You should discuss the approach you have taken to creating the relational database structure, referring to the key aspects of your design, such as why an attribute/variable was selected as a primary key, or why you have elected to contain a specific set of variables/attributes with the same table.

Part 2:

You should discuss and present the following for this task:

* Any further assumptions you made about the scenario, or potential analysis requirements, or reiterate those that are specifically relevant here from Task 1.
* Justification for the technology/technique selected, and the approach to your solution(s), given the nature of the data set and the context of the scenario.
* Clear comparison of benefits and limitations against other potential technology/technique/solutions.
* It is expected that this section will be supported by the literature, with effective use of citations (and attached reference list) to support your claims.

To attain **any Grade** in this task you **MUST** present the requested ERD design and sample SQL in the appendices and use correctly formatted citations and a supporting reference list. Failure to do so will result in a **grade** of **ZERO** for this task.

## Task 3: Considering web-based application (ML05) (20%)

(Suggested word count for this section: 500)

* Clear statement(s) of the three privacy issues you intend to discuss.
* Clear reasoning as to why these are potential issues, and what evidence you have drawn on to identify them.
* Clear presentation of potential mitigations for each of these issues, and where applicable, a comparison with other similar scenarios/issues to support their potential.
* Effective use of citations (and attached reference list) to support your claims.

To attain a **Pass** in this task you **MUST** support your discussion with relevant literature using citations and a reference list, using the IEEE format.

## Referencing

You are required to use the [IEEE referencing style](https://subjectguides.york.ac.uk/referencing-style-guides/ieee) for citing books, articles, and all other sources (such as websites) used in your assignment.

Good referencing is essential in order to meet the standards of academic integrity set by the University. All your sources must be acknowledged, regardless of whether you included direct quotes or not. Visit your **Academic Integrity Tutorial** module in Canvas for additional guidance on effective referencing.

# Marking Criteria

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Learning Outcome** | **Section/Task** | **Criteria** | **Available marks** | |
| **MLO** | **Section/Question** | **Criteria** | | **Marks** |
| **Task 1: Rental “demand” investigation** | | | | |
| 2 | 1 | **Evidence of the work** in Weka and the final data files (in ARFF format only) have been presented. | | **Pass/Fail** |
| 2/3 | 1 | **Approach and results:**  The evidence and discussion present a clear investigation of the data as requested.  Appropriate techniques and pre-processing have been utilised to undertake this.  Results are clear and credible (i.e. not obviously invalid). | | **20** |
| 3 | 1 | **Justification:**  There is a clear and appropriate justification of the approach/techniques used.  This is supported by the literature. | | **10** |
| 3 | 1 | **Critical evaluation:**  The approach has been critically evaluated, which is in alignment with the results generated.  Any identified issues are valid, accurately described, and of genuine concern.  This is supported by the literature. | | **10** |
| **Task 2: Storing data and scalable solutions** | | | | |
| 1 | 2 | **Database design:**  Is correctly presented, and appropriately normalised to 3NF.  The discussion supports and clarifies the approach taken and the decisions made. | | **10** |
| 1 | 2 | **Sample SQL statements**:  Are correct in relation to the presented design. | | **10** |
| 4 | 2 | **Scalable solution:**  A viable approach is described that uses multiple technologies and would be likely to achieve a worthwhile improvement in performance (given coordination overhead etc). | | **20** |
| **Task 3: Considering public-facing application** | | | | |
| 5 | 3 | **Privacy issues:**  Are of genuine concern and are appropriate given the context of the scenario and tasks. | | **10** |
| 5 | 3 | **Mitigation strategies:**  The strategies discussed to deal with the privacy issues offer realistic solutions and are supported by the literature and current standards. | | **10** |
|  | **TOTAL:** | | | **100** |

# Marking Criteria: Grade breakdown

The grade breakdown is written for assessment markers. All statements in each grade band need to be met to be awarded that band.

## Task 1: Rental “demand” investigation – Approach and results – 20%

*MLO2: Manipulate a data set to extract statistics and features.*

*MLO3: Critically evaluate and apply data mining techniques/tools to build a classifier or regression model and predict values for new examples.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * The .ARFF files have not been presented OR the files presented are not in the requested format. * Either no evidence has been presented OR little to none of the presented evidence supports the required areas (discussion, approach and results) effectively. * Either no discussion has been presented OR little to none of the discussion presented for the areas of investigation demonstrates appropriate techniques/approaches, given the nature of the data set and scenario. * Either no results have been presented OR little to none of the results presented for the required areas of investigation are clear or credible, given the nature of the technique/approach used. |
| 40-49% | Marginal fail | * The .ARFF files have been presented and are in the requested format. * Evidence has been presented which supports some of the required areas (discussion, approach or results) to some degree. * The discussion presents some techniques/approaches for the required areas of investigation, some of which are appropriate, but no single area is wholly complete, given the nature of the data set and scenario. * Some results presented for the required areas of investigation are appropriate, but not wholly complete or credible for any single area, given the nature of the technique/approach used. |
| 50%-59% | Pass | * The .ARFF files have been presented and are in the requested format. * Evidence has been presented which supports most of the required areas (discussion, approach or results), with some areas of weakness. * The discussion clearly indicates that at least **one** of the required areas of investigation has utilised an appropriate technique/approach, given the nature of the data set and scenario. * The results for at least **one** of the required areas of investigation are clear and credible, given the nature of the technique/approach used. |
| 60%-69% | Merit | * The .ARFF files have been presented and are in the requested format. * Evidence has been presented which supports all the required areas (discussion, approach or results), with few weak areas. * The discussion clearly indicates that at least **two** of the required areas of investigation have utilised appropriate techniques /approaches, given the nature of the data set and scenario. * The results for at least **two** of the required areas of investigation are clear and credible, given the nature of the technique/approach used. |
| 70%-100% | Distinction | * The .ARFF files have been presented and are in the requested format. * Evidence has been presented which supports all the required areas (discussion, approach or results), with little to no areas of weakness. * All **three** of the required areas of investigation have utilised appropriate techniques/approaches, given the nature of the data set and scenario. * The results for all **three** of the required areas of investigation are clear and credible, given the nature of the technique/approach used. |

## Task 1: Rental “demand” investigation – Justification – 10%

*MLO2: Manipulate a data set to extract statistics and features.*

*MLO3: Critically evaluate and apply data mining techniques/tools to build a classifier or regression model and predict values for new examples.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * The .ARFF files have not been presented OR the files presented are not in the requested format. * Few or no citations or reference list are presented, none of which is in the correct format. * The justification for the techniques/approach used is either not present, incomplete or wholly ineffective/inappropriate for all areas of investigation, given the nature of the data set and scenario. * The justification is not supported by the literature, or the justification is supported in a limited way (many areas of ineffectiveness/inappropriateness). |
| 40-49% | Marginal fail | * The .ARFF files have been presented and are in the requested format. * Some citations and a reference list are presented, but are not necessarily in the correct format. * The justification for the techniques/approach used has some areas of effectiveness/appropriateness, but no single area of investigation is wholly complete, given the nature of the data set and scenario. * The justification is supported by the literature, with a few areas of effectiveness/appropriateness. |
| 50%-59% | Pass | * The .ARFF files have been presented and are in the requested format. * Most citations and the reference list are correctly formatted using IEEE (with minor errors). * There is an effective justification for the techniques/approach used for at least **one** of the investigation areas. * The justification is supported by the literature, with some areas of weakness (in terms of effectiveness/appropriateness). |
| 60%-69% | Merit | * The .ARFF files have been presented and are in the requested format. * Citations and the reference list are correctly formatted using IEEE. * There is a clear and appropriate justification for the techniques/approach used for at least **two** of the investigation areas. * The justification is supported by the literature, with few areas of weakness (effectiveness/appropriateness). |
| 70%-100% | Distinction | * The .ARFF files have been presented and are in the requested format. * Citations and the reference list are correctly formatted using IEEE. * There is a clear and appropriate justification for the techniques/approach used for all **three** of the investigation areas. * The justification is supported by the literature, with few or no areas of weakness (in terms of effectiveness/appropriateness). |

## Task 1: Rental “demand” investigation – Critical Evaluation – 10%

*MLO2: Manipulate a data set to extract statistics and features.*

*MLO3: Critically evaluate and apply data mining techniques/tools to build a classifier or regression model and predict values for new examples.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * The .ARFF files have not been presented OR the files presented are not in the requested format. * Few or no citations or reference list are presented, none of which is in the correct format. * The critical evaluation for the techniques/approach used is either not present, incomplete, misaligned with the presented results, or wholly ineffective/inappropriate for all areas of investigation, given the nature of the data set and scenario. * The critical evaluation is not supported by the literature, or critical evaluation is supported in a limited way (many areas of ineffectiveness/inappropriateness). |
| 40-49% | Marginal fail | * The .ARFF files have been presented and are in the requested format. * Some citations and a reference list are presented but are not necessarily in the correct format. * The critical evaluation for the techniques/approach used has some areas of effectiveness/appropriateness, and some areas may be aligned with the presented results, but no single area of investigation is wholly complete, given the nature of the data set and scenario. * The critical evaluation is supported by the literature, with a few areas of effectiveness/appropriateness. |
| 50%-59% | Pass | * The .ARFF files have been presented and are in the requested format. * Most citations and the reference list are correctly formatted using IEEE (with minor errors). * The critical evaluation provides an accurate and rational appraisal of the effectiveness of the techniques/approach used for at least **one** of the areas of investigation, which is aligned with the presented results. * The critical evaluation is supported by the literature, with some areas of weakness (effectiveness/appropriateness). |
| 60%-69% | Merit | * The .ARFF files have been presented and are in the requested format. * Citations and the reference list are correctly formatted using IEEE. * The critical evaluation provides an accurate and rational appraisal of the effectiveness of the techniques/approach used for at least **two** of the areas of investigation, which is aligned with the presented results. * The critical evaluation is supported by the literature, with few areas of weakness (effectiveness/appropriateness). |
| 70%-100% | Distinction | * The .ARFF files have been presented and are in the requested format. * Citations and the reference list are correctly formatted using IEEE. * The critical evaluation provides an accurate and rational appraisal of the effectiveness of the techniques/approach used for all **three** of the areas of investigation, which is aligned with the presented results. * The critical evaluation is supported by the literature, with few or no areas of weakness (effectiveness/appropriateness). |

## Task 2: Storing data and scalable solutions – Database design – 10%

*MLO1: Create a data set using modern database models and technology.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * The requested ERD design has either not been presented or is not readable. * The ERD diagram demonstrates little or no effective relational database structure (tables and naming are not rational), which demonstrates little or no effective normalisation, and little or no correct notation. * Either no discussion is presented or what is presented describes a limited approach to the decisions made during the construction of the relational database. |
| 40-49% | Marginal fail | * The requested ERD design has been presented, but may not be wholly readable. * The ERD diagram demonstrates a limited relational database structure (tables and naming are not always rational) which has not achieved 3NF, but demonstrates some effective normalisation, and mostly uses correct notation. * The discussion presents the approach to the decisions made during the construction of the relational database, which is rational to some degree. |
| 50%-59% | Pass | * The requested ERD design has been presented and is readable. * The ERD diagram demonstrates a mostly clear and accurate relational database structure (tables and naming are rational) which has been normalised to 3NF, and mostly uses correct notation. * The discussion presents a mostly rational approach to the decisions made during the construction of the relational database. |
| 60%-69% | Merit | * The requested ERD design has been presented and is readable. * The ERD diagram demonstrates a clear and accurate relational database structure (tables and naming are rational) which has been normalised to 3NF, and uses correct notation. * The discussion presents a rational approach to the decisions made during the construction of the relational database. |
| 70%-100% | Distinction | * The requested ERD design has been presented and is readable. * The ERD diagram demonstrates clear and accurate relational database structure (tables and naming are rational) which has been normalised to 3NF, and uses correct notation. * The discussion presents a clear and rational approach to the decisions made during the construction of the relational database. |

## Task 2: Storing data and scalable solutions – Sample SQL statements – 10%

*MLO1: Create a data set using modern database models and technology.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * Either no SQL has been presented, or all presented SQL statements are incorrect and poorly formed to a greater degree (minor and/or major errors across all three statements). |
| 40-49% | Marginal fail | * The presented SQL demonstrates mostly correct (minor errors) for at least one statement, input or output. |
| 50%-59% | Pass | * The presented SQL demonstrates mostly correct (minor errors) for both input and at least one output statement. |
| 60%-69% | Merit | * The presented SQL demonstrates correct construction for both input and at least one output statement. |
| 70%-100% | Distinction | * The presented SQL demonstrates correct construction for both input and all output statements. |

## Task 2: Storing data and scalable solutions – Proposing a scalable solution – 20%

*MLO4: Analyse and communicate issues with scaling up to large data sets and use appropriate techniques to scale up the computation.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * Few or no citations or reference list are presented, none of which is in the correct format. * Either no approach for a scalable solution has been presented or the approach that is presented is unfeasible, with no potential gains identified, and a number of unidentified limitations. * The discussion either presents no justification for the approach, or what is presented is limited, with few or no areas of effectiveness/appropriateness. * The discussion either presents no comparison with alternative technologies/solutions, or what is presented is limited, with few or no areas of effectiveness/appropriateness. * The discussion is not supported by the literature, or the discussion is only supported in a limited way (many areas of ineffectiveness/inappropriateness). |
| 40-49% | Marginal fail | * Some citations and a reference list are presented but are not necessarily in the correct format. * A somewhat viable approach has been presented for a scalable solution, which does not clearly identify any potential gains, or has unidentified limitations. * The discussion presents a somewhat rational justification for the approach, with a few areas of effectiveness/appropriateness. * There is a somewhat clear comparison with some alternative technologies/solutions, which supports the solution to some degree, demonstrating a few areas of effectiveness/appropriateness. * The discussion is supported by the literature, with a few areas of effectiveness/appropriateness. |
| 50%-59% | Pass | * Most citations and the reference list are correctly formatted using IEEE (with minor errors). * A mostly viable approach has been presented for a scalable solution, identifying at least one potential gain. * The discussion presents a mostly clear rationale and justification for the approach, with some areas of weakness (effectiveness/appropriateness). * There is a mostly clear and effective comparison with some alternative technologies/solutions, that mostly supports the solution, with some areas of weakness (effectiveness/appropriateness). * The discussion is supported by the literature, with some areas of weakness (effectiveness/appropriateness). |
| 60%-69% | Merit | * Citations and the reference list are correctly formatted using IEEE. * A viable approach has been presented for a scalable solution, with potential gains and few areas of weakness. * The discussion presents a clear rationale and justification for the approach, with few areas of weakness (effectiveness/appropriateness). * There is a clear and effective comparison with alternative technologies/solutions, that clearly supports the presented solution, with few areas of weakness (effectiveness/appropriateness). * The discussion is supported by the literature, with few areas of weakness (effectiveness/appropriateness). |
| 70%-100% | Distinction | * Citations and the reference list are correctly formatted using IEEE. * A viable and effective approach has been presented for a scalable solution, with clear gains and few or no areas of weakness. * The discussion presents a clear rationale and justification for the approach, with few or no areas of weakness (effectiveness/appropriateness). * There is a comprehensive comparison with alternative technologies/solutions, which clearly supports the presented solution, with few or no areas of weakness (effectiveness/appropriateness). * The discussion is supported by the literature, with few or no areas of weakness (effectiveness/appropriateness). |

## Task 3: Considering public-facing application – Privacy issues – 10%

*MLO5: Critically discuss the need for privacy, identify privacy risks in releasing information, and design techniques to mediate these risks.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * Few or no citations or reference list are presented, none of which is in the correct format. * Either the privacy issue(s) identified are not valid/relevant to the task and scenario’s context, OR they are only discussed in a limited way. * The identified privacy issues are not supported by the literature, or they are only supported in a limited way (many areas of ineffectiveness/inappropriateness). |
| 40-49% | Marginal fail | * Some citations and a reference list are presented but are not necessarily in the correct format. * Although valid privacy issues have been identified that are potentially relevant to the task and scenario, none are discussed in a complete and comprehensive way. * The identified privacy issues are supported by the literature, but with few areas of effectiveness/appropriateness |
| 50%-59% | Pass | * Most citations and the reference list are correctly formatted using IEEE (with minor errors) * At least **one** valid privacy issue has been identified that is relevant to the task and scenario’s context and is discussed in a complete and comprehensive way. * The identified privacy issues are supported by the literature, with some areas of weakness (effectiveness/appropriateness). |
| 60%-69% | Merit | * Citations and the reference list are correctly formatted using IEEE. * At least **two** valid privacy issues have been identified that are relevant to the task and scenario’s context and are discussed in a complete and comprehensive way. * The identified privacy issues are supported by the literature, with few areas of weakness (effectiveness/appropriateness). |
| 70%-100% | Distinction | * Citations and the reference list are correctly formatted using IEEE. * At least **three** valid privacy issues have been identified that are relevant to the task and scenario’s context and are discussed in a complete and comprehensive way. * The identified privacy issues are supported by the literature, with few or no areas of weakness (effectiveness/appropriateness). |

## Task 3: Considering public-facing application – Mitigation strategies – 10%

*MLO5: Critically discuss the need for privacy, identify privacy risks in releasing information, and design techniques to mediate these risks.*

|  |  |  |
| --- | --- | --- |
| 0-39% | Fail | * Few or no citations or reference list are presented, none of which is in the correct format. * Either no mitigation strategies at all are presented, OR if they are, they are not feasible, OR do not clearly align with any given privacy issues, OR are not relevant given the context of the task and scenario. * The mitigation strategies are not supported by the literature, or they are only supported in a limited way (many areas of ineffectiveness/inappropriateness). |
| 40-49% | Marginal fail | * Some citations and a reference list are presented but are not necessarily in the correct format. * Mitigation strategies are presented, some of which are feasible, but they either do not clearly align with given privacy issues, or the context of the task and scenario. * The mitigation strategies are supported by the literature, but with few areas of effectiveness/appropriateness |
| 50%-59% | Pass | * Most citations and the reference list are correctly formatted using IEEE (with minor errors) * At least **one** of the mitigation strategies presented is feasible, and aligns clearly with relevant given privacy issues, and the context of the task and scenario. * The mitigation strategies are supported by the literature, with some areas of weakness (effectiveness/appropriateness). |
| 60%-69% | Merit | * Citations and the reference list are correctly formatted using IEEE. * At least **two** of the mitigation strategies presented are feasible, and align clearly with relevant given privacy issues, and the context of the task and scenario. * The mitigation strategies are supported by the literature, with few areas of weakness (effectiveness/appropriateness). |
| 70%-100% | Distinction | * Citations and the reference list are correctly formatted using IEEE. * At least **three** of the mitigation strategies presented are feasible, and clearly align with relevant given privacy issues, and the context of the task and scenario. * The mitigation strategies are supported by the literature, with few or no areas of weakness, (effectiveness/appropriateness). |

# Assessment Submission

You will submit your assessment in the ‘Assignments’ area of the module in Canvas. Please check your Canvas module for the specific submission date for this assignment.

We recommend that you allow at least 30 minutes before the deadline to upload your submission, as failure to upload your assessment file within the allotted time is not admissible as an exceptional circumstance.

The webpage [How do I submit an online assignment?](https://community.canvaslms.com/t5/Student-Guide/How-do-I-submit-an-online-assignment/ta-p/503) provides further technical information on how to upload an assessment. The advice given here comes directly from Canvas. We do not recommend uploading assignments by mobile. We recommend you view the submission after uploading your work to ensure the correct file has been submitted and no technical errors have occurred.

If you face any technical difficulties whilst trying to submit this assessment, then contact Canvas support on [support@instructure.com](mailto:support@instructure.com) or +44 80 0060 8442 (available 24 hours) in advance of the deadline. You should also email [cs-online-admin@york.ac.uk](mailto:cs-online-admin@york.ac.uk) as a matter of urgency to report the issue and receive further instruction.

# Assessment Policies

This assessment is subject to the policies stated on the ‘Summative Assessment Policies’ page in Canvas. These policies include (but are not limited to):

* Academic Integrity and submission of student work to Turnitin
* Advice on anonymising your assessment
* Penalties for late submission
* Marking policy for multiple submissions
* The Fit to Sit / Submit policy
* Passing mark and module reassessment

Please ensure that you have read and understood these policies before starting the assessment.