



COM6013

Data Mining

Portfolio

Date for Submission: Please refer to the timetable on ilearn

**(The submission portal on ilearn will close at 14:00 UK time
on the date of submission)**

Page 1 of 10

[2874]

Arden University © reserves all rights of copyright and all other intellectual property rights in the learning materials and this publication. No part of any of the learning materials or this publication may be reproduced, shared (including in private social media groups), stored in a retrieval system or transmitted in any form or means, including without limitation electronic, mechanical, photocopying, recording or otherwise, without the prior written consent of Arden University. To find out more about the use and distribution of programme materials please see the Arden Student Terms and Conditions.

Template: V5



Assignment Brief

As part of the formal assessment for the programme you are required to submit a **Data Mining** assignment. Please refer to your Student Handbook for full details of the programme assessment scheme and general information on preparing and submitting assignments. The assignment brief will specifically give details and instructions for the assignment. No examination, or details of, are included within this module.

Module description: Coursework 100%

Description: The assignment is given as two tasks. Each task is to be answered individually, unless multiple components have been given. Task 1 has been designed to check your understanding of Data Mining methodologies for critical analysis, through a real-world mini Data Mining project. All of which will be written in Python.

Task 2 is designed to give you the opportunity to demonstrate your Data Mining reporting skills, comprehensively written in an easily understandable manner, based upon your Data Mining study.

A clear, concise analysis of the dataset for all Tasks are to be given within the submission, complimented with screenshot evidence of all processes and results. You are to submit a single Word document for all tasks and your code for Task 1 is to be included within an appendix, so that it can be checked and verified it is working correctly. Your student ID number must be clearly defined upon the uploaded file.



Learning Outcomes:

	Coursework
Understand the theoretical concepts and motivation behind data mining in the context of business and information technology. LO 1	X
Evaluate the strengths and weakness of different data mining techniques for a given task. LO 2	X
Apply and where necessary, adapt an appropriate data mining technique to a given problem. LO 3	X
Graduate Attribute Reflective Practitioner: Undertake critical analysis and reach reasoned and evidenced decisions, contribute problem-solving skills to find innovate in solutions. LO 4	X

All learning outcomes must be met to pass the module.



Guidance

Your assignment should include: a title page containing your student number, the module name, the submission deadline and the exact word count of your submitted document; the appendices if relevant; and a reference list in (see referencing section for more information). You should address all the elements of the assignment task listed below. Please note that tutors will use the assessment criteria set out below in assessing your work.

You must not include your name in your submission because Arden University operates anonymous marking, which means that markers should not be aware of the identity of the student. However, please do not forget to include your STU number.

Maximum word count: 3000 words

Please refer to the full word count policy which can be found in the Student Policies section here: [Arden University | Regulatory Framework](#)

The word count includes everything in the main body of the assessment (including in text citations and references). The word count excludes **numerical data in tables, figures, diagrams, footnotes, reference list and appendices. All other printed words ARE included in the word count.**

Students who exceed the wordcount up to a 10% margin will not be penalised. Students should note that no marks will be assigned to work exceeding the specified limit once the maximum assessment size limit has been reached.



Assignment Task

Task 1 – Data Mining Mini Project in Python

Detail Task 1

You have recently secured as position as a junior data analyst at a Marine Biology data investigation company. As your first data mining investigation, your line manager has asked you to build and complete appropriate models on the health of the Abalone populations within the seas surrounding the West coast of Australia. In doing so, the company can use your optimum, relative comparative investigation to determine the health of the Abalone's population, within the geographical area, for Males (M), Female (F) and Infants (I) based upon differing variables.

Before the analysis can be conducted, ensure that all pre-processing has been undertaken on the data set and the data set is fully understood through initial descriptive statistics. For the main body of the analysis, you are to choose the most appropriate comparative modelling techniques and processes for the investigation, based upon the data. In doing so, this will allow insight into the overall analysis and appropriate meaningful solutions. The full investigation is to be conducted in Python.

(70 Total marks)

(LO's: 2 & 3)

(2000 Word Equivalence for the code)

Task 2 – Data mining report

Detail task 2

Based upon the completed data mining investigation in Task 1, you will now write a comprehensive report that can be easily understood by the company's board of directors and higher management. Contained within you will include full reasoning for your chosen Modelling techniques and solutions, in addition to your motivation, investigations processes, results and conclusions.

(30 Total marks)

(LO's: 1 & 4)

End of questions



As technology and platforms may change, your module tutor will provide you with up-to-date details.

Formative Feedback

to receive formative feedback

You are encouraged to submit your assignment for feedback **once** and it is 30% of your entire submission. You, the student, are to choose 30%, **not the tutor**. The last day for guaranteed feedback hand in is Monday of week 9 at 23:58. No formative feedback will be given after the time specified above for blended learning.

The Feedback is designed to help you develop areas of your work, encouraging academic skills and independent learning.

If you are a Distance Learning student, then you are encouraged to send 30% of your assignment for feedback by email to your tutor, no later than two weeks before your final submission date. Dates will be given to you by your tutor on a module by module basis.

Referencing Guidance

You **MUST** underpin your analysis and evaluation of the key issues with appropriate and wide ranging academic research and ensure this is referenced using the AU Harvard system(s).

Follow this link to find the referencing guides for your subject: [Arden Library](#)

Submission Guidance

Assignments submitted late will not be accepted and will be marked as a 0% fail.

Your assessment can be submitted as a single Word (MS Word) or PDF file, or, as multiple files.

If you chose to submit multiple files, you must name each document as the question/part you are answering along with your student number ie Q1 Section A STUXXXX. **If you wish to overwrite your submission or one of your submissions, you must ensure that your new submission is named exactly the same as the previous in order for the system to overwrite it.**

You must ensure that the submitted assignment is all your own work and that all sources used are correctly attributed. Penalties apply to assignments which show evidence of academic unfair practice. (See the Student Handbook which is available on the A-Z key information on iLearn.)



Assessment Criteria (Learning objectives covered - all)

Level 6 study represents the student's increasing autonomy and independence in relation to their knowledge, understanding and skills. At Level 6, students are expected to demonstrate problem solving skills in both practical and theoretical contexts. This should be supported by an understanding of appropriate theory, creativity in expression and thought based on independent but informed judgments. Students should demonstrate the ability to seek out, invoke, analyse, and evaluate competing theories and claims to knowledge and work in a critically constructive manner. Work at this level is articulate, coherent, and skilled.

Grade	Mark Bands	Generic Assessment Criteria
First (1)	80%+	<p>An outstanding knowledge base exploring and analysing the discipline, its theory, and any associated ethical considerations. The work demonstrates outstanding independence of thought and originality. There is outstanding management of learning resources, and a high degree of autonomy is demonstrated which goes above and beyond the brief.</p> <p>The work demonstrates intellectual originality and creativity. Writing is exceptionally well structured and accurately referenced throughout. Where appropriate, outstanding professional skills are demonstrated. The work is original and with some additional effort could be considered for internal publication.</p>
	70-79%	<p>An excellent information base within which the discipline is explored and analysed.</p> <p>There is considerable originality in the approach and the work demonstrates confidence and autonomy and extends to consider ethical issues. Learning resources have been managed with exceptional confidence and the work exceeds the assessment brief. Writing is exceptionally well structured and accurately referenced throughout.</p> <p>Where appropriate, an excellent level of professional skills are demonstrated and the work demonstrates a high level of intellectual and academic skills.</p>
Upper second (2:1)	60-69%	<p>A very good knowledge base which explores and analyses the discipline, its theory, and any associated ethical issues.</p> <p>There is evidence of some originality and independence of thought. A very good range of learning resources underpin the work and there is clear evidence of self-directed research. The work demonstrates the ability to analyse the subject and apply theory with good academic and intellectual skills. Academic writing skills are very good, expression is accurate overall, and the work is consistently referenced throughout.</p>



Lower second (2:2)	50-59%	A good understanding of the discipline which supports some analysis, evaluation and problem-solving within the discipline. There may be reference to some of the ethical considerations. The work shows a sound level of competence in managing basic sources and materials. Academic writing skills are good and accurate overall, and the work is planned and structured with some thought. Professional skills are good (where appropriate). The work may lack originality, but academic and intellectual skills are moving into the critical domain. The work is referenced throughout.
Third (3)	40-49%	Satisfactory level of performance in which there are some omissions in the understanding of the subject, its underpinning theory, and ethical considerations. There is little evidence of independent thought, and the work shows a basic use of sources and materials. Academic and intellectual skills are limited. The work may lack structure overall. There are some difficulties in developing professional skills (where appropriate). There is an attempt to reference the work.
Marginal Fail	30-39%	A limited piece of work in which there are clear gaps in understanding the subject, its underpinning theory, and ethical considerations. The work shows a limited use of sources and materials. Academic and intellectual skills are weak and there are errors in expression and the work may lack structure overall. There are difficulties in developing professional skills (where appropriate). The work lacks original thought and is largely imitative.
Clear fail	29% and Below	A poor performance in which there are substantial gaps in knowledge and understanding, underpinning theory and ethical considerations. The work shows little evidence in the use of appropriate sources and materials. Academic writing skills are very weak and there are numerous errors in expression. The work lacks structure overall. Professional skills (where appropriate) are not developed. The work is imitative.



Rubric Task 1: Data Mining Project (70 Marks)

Criteria	Outstanding 80% - 100%	Excellent 70% - 79%	Very Good 60% - 69%	Good 50% - 59%	Pass 40% - 49%	Limited 0 – 39%
Task 1 Data Mining Project (70% Marks Overall)	An outstanding understanding of the processes involved when conducting a Data Mining investigation. Data importing, pre-processing and descriptive statistics have all be completed correctly. The correct data mining methodology has been Applied. All processes and solutions are correct.	An excellent understanding of the processes involved when conducting a Data Mining investigation. Data importing, pre-processing and descriptive statistics have all be completed. The data mining methodology has been Applied. Processes are correct.	A very good understanding of most of the processes involved when conducting a Data Mining investigation. Data importing, pre-processing and descriptive statistics have all be completed. The data mining methodology has been Applied. Processes are mostly correct and justified.	A good understanding of most processes involved when conducting a Data Mining investigation. Some Data importing, pre-processing and descriptive statistics have all be completed. The data mining methodology has been mostly applied. Some processes are correct and justified.	A basic understanding of the processes involved when conducting a Data Mining investigation. Minimal Data importing, pre-processing and descriptive statistics have all be completed. Minimal data mining methodologies have been applied. Processes and solutions have been given, yet incorrect.	A limited understanding of the processes involved when conducting a Data Mining investigation. No pre-processing and statistical investigation has been completed. No data mining methodology has been applied and no processes are evident

Rubric Task 2: Comprehensive Report (30 Marks)

Criteria	Outstanding 80% - 100%	Excellent 70% - 79%	Very Good 60% - 69%	Good 50% - 59%	Pass 40% - 49%	Limited 0 – 39%
Task 2 Report writing (30% Marks Overall)	An outstanding report. All analysis results have been given and are correct, therefore inferring the correct data mining methodologies. All results are fully reasoned justified. No gaps are shown within the analysis.	An excellent report. Analysis results have been given and correct, therefore inferring the appropriate data mining methodologies. Results are reasoned and justified.	A very good report. Most Analysis results have been given, therefore mostly inferring some form of data mining methodology. Most of the Results are reasoned with and justified. Minor gaps are shown within the analysis.	A good report. Some Analysis results have been given, therefore inferring some form of data mining methodology. Some of the Results are reasoned with and justified. Gaps are shown within some of the analysis.	A basic report. Some Analysis results have been given, yet incorrect. Some of the Results are reasoned with and justified, but incorrect. Frequent gaps are shown within the analysis.	A poor report. No Analysis has been given. No Results are reasoned with or justified. Major Gaps throughout the analysis.



Note: All sub-tasks are considered to be equally weighted. Where grades are awarded across multiple sub-tasks, the marker will allocate grades appropriately.

e.g. If the task contains a part a) and b) they are considered together (as each relates to different aspects of the learning outcomes) and, a poor response to the second part would reduce the overall grade awarded, even if the response to the first part is of a high standard,