Middlesex University London

Faculty of Science and Technology

CST3340

Business Intelligence

Module Leader: Joanna Loveday (Hendon)

Term (AY) 2020-21

Duration of the module in weeks: 24 Weeks

Document Version V1.

Online location of handbook

This handbook can be accessed via MyLearning.

Other formats available

This handbook is available in a large print format. If you would like a large print copy or have other requirements for the handbook, please contact the Disability Support Service (disability@mdx.ac.uk, +44 (0)20 8411 4945).

Disclaimer

The material in this handbook is as accurate as possible at the date of production. You will be notified of any minor changes. If there are any major changes to the module you will be consulted prior to the changes being confirmed. Please check the version number on the front page of this handbook to ensure that you are using the most accurate information.

Other documents

Your module handbook should be read and used alongside your programme handbook and the information available to all students on My Learning, including the Academic Regulations. Your programme handbook can be found on the My Learning programme page.

Table of Contents

1. Module Introduction

Large scale data management is now the central focus of most industrial data processing applications and the driving force behind Business Intelligence. This module covers the concepts and theories of managing very large data, namely data warehouse and Big Data and deals with the concepts and theories of modelling of large volume data storage and preserving data quality. The module also explores technologies and algorithms for handling a large amount of data, i.e., data cleansing, data segmentation and data transformation, to discover knowledge from large data sources.

The module provides students with an opportunity to explore, review and examine how large-scale data can be managed in industry to enhance the decision-making power of managers. The module also equips students with sufficient understanding of the current theory of decision making and its applications that address modern and future business needs using business intelligence.

2. The Module Team

Joanna Loveday	
Role	Module Tutor
Room number:	T131 [Town Hall Building]
Email:	j.loveday@mdx.ac.uk
Telephone number:	02084116613
Office Hours:	TBC

3. Staff Student Communication

Students may contact staff via e-mail, phone, by dropping in to staff office hours, and by making an appointment to see them outside office hours.

Staff will contact students by e-mail, phone, the My Learning module page and via lectures and seminars.

The team may send urgent group and/or individual messages about the module to you by email, so it is important that you read your University email regularly.

All staff have office hours, it is not necessary to book an appointment during these hours, you just need to drop-in.

In the first instance problems should be dealt with by talking to a member of the module team. You can give feedback on this module to the module leader, your Student Voice Leader, to your personal tutor, and through the end of module evaluation survey.

Our most important consideration is your health, wellbeing and safety as well as our staff and people related to the University. Remember that you – as part of #TeamMDX – can stay up-to-date with the guidance on Coronavirus at

https://unihub.mdx.ac.uk/coronavirus-covid19

4. Module Overview

Aims

This module aims to develop an understanding of the techniques and approaches used to capture, store and analyse data generated by organisations for purposes of business intelligence. In a digital age it is important for businesses to make use of data captured about its entities. You will learn about information retrieval, data presentation, pattern recognition techniques and data models that can be used in business intelligence applications. Descriptive data models can be used to gain a better understanding of overall organisation and predictive data models inform the decision making in all aspects of the business.

Learning Outcomes

Knowledge

On completion of this module, the successful student will be able to:

1. Analyse the data architecture, business processes and the process of establishing data relationship in order to recommend relevant data stores that would assist decision making.
2. Critically evaluate different business intelligence concepts and techniques to recommend an appropriate approach for incorporating business intelligence in the business decision making process.
3. Critically evaluate current business intelligence tools suitable for data analysis and mining in a range of business application areas
4. Demonstrate knowledge and understanding of the professional, legal, social and ethical issues related to the use of business intelligence technologies.

Skills

This module will call for the successful student to demonstrate:

5. The ability to apply a variety of architectures and technologies for data collection, integration and storage.
6. The ability to identify relevant data intelligence techniques for addressing a variety of business decision problems.
7. The ability to apply contemporary business intelligence tools for information retrieval, data presentation, data visualisation and pattern recognition.
8. The ability to professionally present results from the business decision making process, verbally and in writing.

Syllabus

- Data Warehouse Concepts and Architecture
- Multidimensional Data Modelling and Online Analytical Processing (OLAP)
- Data integration and Extract, Transform, Load Process (ETL)
- Big Data and Alternate Technologies e.g. Hadoop, MapReduce and NoSQL.
- Business Intelligence Concepts and Techniques
- Data/Information Visualisation
- Data Mining Concepts and Algorithms
- Selected professional, ethical and legal issues, including GDPR

Learning hours

- Scheduled teaching 3 hours per week
- Independent study 10 hours per week

Research Ethics –

- **The teaching, learning, assessment and research activities undertaken in this module have been considered and are <u>not likely to require ethical approval</u>.**
- However, please seek advice if undertaking the module entails carrying out any research activities involving **human participants, human data, animals/animal products, precious artefacts, materials or data systems.** If you submit work that includes data gathered from or about people, this may be treated as academic misconduct and could lead to fail grade being awarded.
- Research ethics approval seeks to ensure all research is designed and undertaken according to certain principles of ethical research. These include:

1. Primary concern must be given to the **safety, welfare and dignity** of participants, researchers, colleagues, the environment and the wider community
2. Consideration of **risks** should be undertaken before research commences with the aim of minimising risks to those involved – i.e. human participants or animal subjects, colleagues, the environment and the wider community, as well as actual or potential risks to those directly or indirectly affected by the research.
3. **Informed consent** should be freely given by participants, and by a trained person when collecting or analysing human tissue (details on accessing and completing online training for gaining informed consent for HTA purposes can be found below in Section 8).
4. Respect for the **privacy, confidentiality and anonymity** of participants
5. Consideration of the rights of **people who may be vulnerable** (by virtue of perceived or actual differences in their social status, ethnic origin, gender, mental capacities, or other such characteristics) who may be less competent or able to refuse to give consent to participate
6. Researchers have a responsibility to the general public and to their profession; as such they should balance the anticipated benefits of their research against **potential harm, misuse or abuse** which must be avoided
7. Researchers must demonstrate the highest standards of **ethical conduct and research integrity**. They must work within the limits of their skills, training and experience, and refrain from exploitation, dishonesty, plagiarism, infringement of intellectual property rights and the fabrication of research results. They should declare any actual or potential conflicts of interest, and where necessary take steps to resolve them.
8. When using human tissues for research, Human Tissue Act and Human Tissue Authority (HTA) requirements must be met. Please contact the relevant designated person (DP) in your department or the HTA Designated Individual (DI) (Dr Lucy Ghali - L.Ghali@mdx.ac.uk). Further information is provided below in the section: "Human Tissue Authority Information", see 'Governance Structure" document and SOPs etc.
9. Research should **not involve any illegal activity**, and researchers must comply with all relevant laws.

- For more information about ethics go to the Middlesex Online Research Ethics (MORE) system which has information and guidance to help you meet the highest standards of ethical research using this link: https://MOREform.mdx.ac.uk

- Information and further guidance on how to complete a research ethics application form (e.g., video guides and templates) can be found on the MORE MyLearning site*: http://mdx.mrooms.net/enrol/index.php?id=12277 (Log in required)

*Middlesex University Definition of Research document can be located on this site.

5. Learning Resources

This module has a variety of learning resources available for you to use to support your learning. These include module notes, worked examples, solutions to exercises, feedback, podcasts, and key reading materials. These can be accessed online via the module page. Please visit the module page regularly to make use of these.

6. Making the most of this module

The module team are here to help and support you achieve your goals. One of the key elements to successfully completing this module is engaging with all of the learning opportunities we offer as well and working with your peers to support one another.

Engaging with online and remote learning and activities is integral to your success. Middlesex University supports students, enabling them to achieve their full potential.

We provide this support through a number of strategies, all of which provide our students with a supportive learning environment online, remotely, face-to-face, or blended.

Further information on engaging with your programme will be available at your Induction and updates online at UniHub

https://unihub.mdx.ac.uk/study/assessment/attendance

Participation and engagement

This module is designed as a combination of contact sessions, directed study and independent study. This means you must participate in all the allocated sessions and you must complete all set prework and activities outside them. Students are expected to take an active part in all learning sessions whether these are online or on campus; lectures, lab sessions, practical classes, seminars and workshops.

To make the most of this module please complete the following every week

- Complete all prework in preparation for learning sessions. This may be watching videos, reading through set material or chapters and completing

activities. Please make notes of points you need to clarify and discuss these in learning sessions with module tutors.
- Read through the notes making a note of any points you need to discuss with your tutor.
- Complete the set activities before the next session, making a note of any points you need to discuss with your tutor.
- Go to the module My Learning page, attempt the quizzes, make use of extra material, view the podcasts, and access the activity solutions. Make a note of anything you wish to discuss with your tutor.
- Complete further reading from the core text online.

The module team is committed to support you and your fellow students whilst you undertake this module. In order for you to get the most out of sessions you need to come prepared and ready to contribute. Please ensure that any work set by the team has been completed before workshops. After each class please review what has been covered and make a note of anything you would like clarification on.

It is important that you are respectful and supportive to your fellow students and tutors. Adopting this approach will create a positive atmosphere within sessions and is something you can use in your professional life.

To access some of the rooms and specialist space used for this module you will need your University ID card. Please remember that your University ID should be carried with you always.

Academic Integrity and misconduct

Academic misconduct is a breach of the values of academic integrity and can occur when a student cheats in an assessment or attempts to deliberately mislead an examiner that the work presented is their own when it is not. It includes, but is not limited to, plagiarism, commissioning or buying work from a third party or copying the work of others, breach of examination room rules.

Students who attempt to gain unfair advantage over others through academic misconduct will be penalised by sanctions, according to the severity of the offence, which can include exclusion from the University. Links to the relevant University regulations and additional support resources can be found here:

Becoming a successful student Course which includes Academic Integrity **Access to course**. (You will have to log into to MyUniHub and then MyLearning to access the course)
Section F: Infringement of Assessment Regulations/Academic Misconduct:
https://www.mdx.ac.uk/about-us/policies/university-regulations
Referencing & Plagiarism: Suspected of plagiarism?
http://libguides.mdx.ac.uk/c.php?g=322119&p=2155601
Referencing and avoiding plagiarism:
http://unihub.mdx.ac.uk/your-study/learning-enhancement-team/online-resources/referencing-and-avoiding-plagiarism

The MDXSU Advice Service offers free and independent support face-to-face in making an appeal, complaint or responding to any allegations of academic or non-academic misconduct.
https://www.mdxsu.com/advice

Extenuating circumstances:
There may be difficult circumstances in your life that affect your ability to meet an assessment deadline or affect your performance in an assessment. These are known as extenuating circumstances or 'ECs'. Extenuating circumstances are exceptional, seriously adverse and outside of your control. Please see link for further information and guidelines:

https://unihub.mdx.ac.uk/your-study/assessment-and-regulations/extenuating-circumstances

7. Module overview and learning schedule

All sessions will be online

Week beginning	Week (s)	Lecture	Lab	Staff	Student Activity	Assessment and feedback
Sept 28th	1	Introduction to Business Intelligence	What is Business Intelligence? Discussion of examples of Business Intelligence usage Data requirements for Business Intelligence Structured Data vs Big Data	JL	Complete lab exercises for Week 1	
Oct 5th to Oct 26th	2 - 5	Introduction to Data Warehousing Multidimensional Data Modelling Data Warehouse Schemas ETL Process	Introduction to Data warehouse; What Is a Data Warehouse? Differences between Operational Database Systems and Data Warehouses Multidimensional Data Model; Data Cubes Stars, Snowflakes, and Fact Constellations: Schemas for Multidimensional Databases; Examples	JL	Complete lab exercises for weeks 2 to 5	Working on Coursework 1

			Types of OLAP; Servers: ROLAP versus MOLAP versus HOLAP ETL Process; Examples			
Nov 2nd to Nov 9th	6 - 7	Alternate data storage i.e. Hadoop	Introduction and discussion of alternative data storage for large data sets.	JL	Complete lab exercises for weeks 6 to 7	Working on Coursework 1
Nov 16th	8	Coursework 1 Formative Feedback		JL		Formative Feedback given during lab sessions
Nov 23rd	9	Data Ethics	Selected professional, ethical and legal issues, including GDPR	JL	Complete lab exercises for week 9	Working on Coursework 1
Nov 30th	10	Coursework 1 Submission week		JL		Formative Feedback given during lab sessions; Final submission midnight on Friday
Dec 7th	11	Introduction to Business Intelligence and Visualisation	Introduction to Business Intelligence and Visualisation using Tableau;	JL	Complete lab exercises for week 11	

Jan 16th to Feb 1st	12-16	Business Intelligence/ Visualisation	Analysing/visualising data using BI tools such as Tableau	JL	Complete lab exercises for weeks 12 to 16	Working on Coursework 2
Feb 8th to Mar 15th	17-22	Data Mining including Data pre-processing; Association and Correlations; Classification and Prediction; Cluster Analysis	**Data pre-processing;** **Association Rules:** Mining Frequent Patterns; Market Basket Analysis: A Motivating Example; Frequent Item-sets and Association Rules; The Apriori Algorithm: Finding Frequent Item-sets; Mining Various Kinds of Association Rules; **Classification and Prediction:** Decision Tree Induction - Tree Pruning Bayesian Classification Neural Networks Regression Time Series Analysis **Cluster Analysis**: What Is Cluster Analysis? K-Means	JL	Complete lab exercises for weeks 17 to 22	Working on Coursework 2

Mar 22nd	23	Coursework 2 Formative Feedback		JL		Formative Feedback given during lab sessions
Apr 12th	24	Coursework 2 Submission Week		JL		Formative Feedback given during lab sessions; Final submission midnight on Friday

8. Assessment

Formative assessment: Formative assessments do not directly contribute to the overall module mark but they do provide an important opportunity to receive feedback on your learning. They provide an opportunity to evaluate and reflect on your understanding of what you have learnt. They also help your tutors identify what further support and guidance can be given to improve your grade.

On this module you will have the following format assessment opportunities.

- Coursework 1 – Feedback week in week 8
- Coursework 2 – Feedback week in week 23

During feedback week, a draft submission will be presented to your tutor at the seminar/lab session and feedback given. At this stage the work will not be marked, just checked for completeness.

Summative assessment: Summative assessment is the assessed work that determines the overall module grade. It is the way the University verifies that students have met the learning outcomes at the appropriate level.

There are 3 assessment components in this module, see assessment scheme below:

During submission week, the final version of the coursework should be submitted via the submission link on unihub.

Students need to pass all the coursework components to achieve a pass in the module overall.

Assessment	Type	Task	Method of submission	Submission week	% of the module overall grade
Coursework 1	Individual Report	Design of data strategy for a chosen scenario. 2400 word report on the data strategy design and implementation for the chosen scenario.	Electronic copy presented during seminar session Online submission through Unihub	Formative Feedback: Week 8 Week 10	Report: 40%
Coursework 1		Summative Feedback	Online via Unihub	Week 13	
Coursework 2	Individual Report Oral Presentation	Data analysis including data visualisation, data pre-processing and data mining of a chosen data set. 2400 word report on the data analysis carried out 5 minute online oral presentation to discuss the data analysis results	Electronic copy presented during seminar session Online submission through Unihub Online in allotted time slot	Formative Feedback: Week 23 Report: Week 24 Oral presentation: During allocated time slot.	Report: 40% Oral Presentation: 20%
Coursework 2		Summative Feedback	Online via Unihub	June 2021	
					100%

Coursework 1

a. Overview

This section contains the material for coursework 1. It is worth 40% of the final grade for this module.

b. Important Dates

You must ensure you meet the deadlines for each element.

Element	Type	Due Date
Individual CW1 Report	Formative Feedback	Labs in Week 8: Week begining: 16th November 2020
	Final Report	Week 10: Midnight Friday 4th December 2020
Summative Feedback		Week 13: Friday 15th January 2021

c. Plagiarism

Plagiarism in data warehouse design is easy to spot. Please be aware that the penalties are severe, and your final degree award classification is at risk.

d. Coursework Content

Coursework 1 should be done individually.

e. Where to submit

Each student should submit an individual report of no more than **2400-word**. Please do not include your name or student Id as your submission will be marked anonymously.

All submission should be via the submission link in your learning environment.

Do not handwritten assessed coursework directly to your tutor, and do not submit it by email to your tutor. Coursework which is not submitted via unihub will not be accepted.

f. Coursework Requirements:

This coursework is based on a case study of your choice. The case study should represent a scenario where data can be used to improve the decision-making process. The submission should be in the form of a report of no more than **2400 word**, see below for more details. Your report will be based on the material covered in the lectures and labs, but you are also expected to carry out your own research.

Your submission should contain the following sections:

- The Case Study (15 marks)
  - A brief overview of your chosen case study, which clearly describes the scenario where data can be used to improve the decision-making process.

- The Data Requirements (15 marks)
  - A detailed discussion of the data requirements of your scenario. This should include a discussion of both:
    - Internal data, data collected locally.
    - External data, data collected by external bodies.

- The Data Warehouse (30 marks)
  - A discussion of data identified which would be suitable for storage in a data warehouse.
  - An example and brief discussion of one suitable data schema. E.g. A star, snowflake or galaxy schema.
  - Using examples from your scenario, discuss:
    - The ETL process.
    - Online Analytical Processing (OLAP).

- Big Data (25 marks)
  - A discussion of data identified which would **not** be suitable for storage in a data warehouse.
  - A discussion of reasons the data is **not** suitable for storge in a data warehouse.
  - Using examples from your scenario, discuss a framework that could be used to collect, store and analyse this data.

- Conclusion (15 marks)
  - This should include a summary of the report and suggest the most appropriate data strategy for your scenario.

Coursework 2

a. Overview

This section contains the material for coursework 2. It is worth 60% of the final grade for this module.

b. Important Dates

You must ensure you meet the deadlines for each element.

Element	Type	Due Date
Individual CW2 Report 40% of the final module grade	Formative Feedback	During Labs in Week 23: Week beginning: 22nd March 2021
	Final Report	Week 24: Midnight Friday 16th April 2021
Oral Presentation for CW2 20% of the final module grade	Presentation	During allocated time slot: Week beginning 19th April 2021
Summative Feedback		June 2021

c. Plagiarism

Plagiarism in data analysis is easy to spot. Please be aware that the penalties are severe, and your final degree award classification is at risk.

d. Coursework Content

Coursework 2 should be done individually.

e. Where to submit

Each student should submit an individual report of no more than **2400-word**. Please do not include your name or student Id as your submission will be marked anonymously. You are also expected to attend a **5-minute** oral presentation**. The presentation is compulsory, students who do not attend the presentation will fail this coursework.**

All submission should be via the submission link in your learning environment.

Do not handwritten assessed coursework directly to your tutor, and do not submit it by email to your tutor. Coursework which is not submitted via unihub will not be accepted.

f. Coursework Requirements:

You are required to analyse a large data set of your choice, which has been agreed with your module tutor:

Your project may use any combination of data analysis techniques, data-mining algorithms and software that has been covered in the module. You may also apply them to any aspect(s) of the dataset for knowledge discovery.

You should cover the areas indicated below and your findings should be presented in the form of a report no more than **2400 word**. You will also be expected to give a **5-minute** oral presentation. **The presentation is compulsory, students who do not attend will fail this coursework.**

Please see below the aspects that you should consider:

Individual Report

- Data Analysis and Visualisation (35 marks)
  - Initial analysis of the data using visualisation techniques within Tableau (use diagrams/graphs to highlight important patterns/findings).
  - Discussion and interpretation of result.
  - Discussion of overall trends and patterns observed.

- Selection of Data Mining Algorithm (10 marks)
  - Select one data mining algorithm suitable for further analysis of your data.
  - Clearly justify your choice, with reference to the visualisation analysis carried out.

- Data Pre-processing (10 marks)
  - Identify your input and class variables, if relevant (i.e. which variable are you going to consider for your class variables).
  - Identify and resolve any anomalies in the data (i.e. missing values, outliers etc.).
  - Carry out any appropriate pre-processing/transformations to the data set.

- Data Mining (25 marks)
  - Use the chosen data mining algorithm for further analysis of your pre-processed data set.
  - Clearly discuss the implementation of the data mining algorithm.
  - Discuss and interpret the results.

- Data Ethics (10 marks)
  - A discussion of data ethical issues related to the analysis and use of business data.

- Conclusion (10 marks)
  - A discussion of the overall visualisation results (e.g. What were the important findings? Summary of overall trends and patterns).
  - A discussion of the data mining results (e.g. How well did the model fit your data?).
  - A discussion of the business intelligence that can be obtained from these results.

- **Oral Presentation** (100 marks)
  - This **5-minute** oral presentation will allow you to discuss your analysis and results.

Feedback on your assignments

You will be provided with feedback on all assessment that is helpful and informative, consistent with aiding the learning and development process.

Feedback will normally be provided within 15 working days of the published assessment component submission date.

Overall module grade

Each component of assessment will be marked as a percentage. To produce the overall module grade a weighted average percentage will be calculated and then converted to a 20-point grade using the scale below.

In order to pass this module, you need to pass all assessment tasks with a minimum grade of *35% with an overall module grade of 40%*

Before you submit your work for final grading, please ensure that you have accurately referenced the work. It is your responsibility to check the spelling and grammar. If you have submitted a formative or draft assessment, you will receive feedback but no grade. The comments should inform you about how well you have done or tell you about the areas for improvement. All assignments should be submitted online unless specified in assessment briefs.

Reassessment for this module normally takes place during the summer semester.

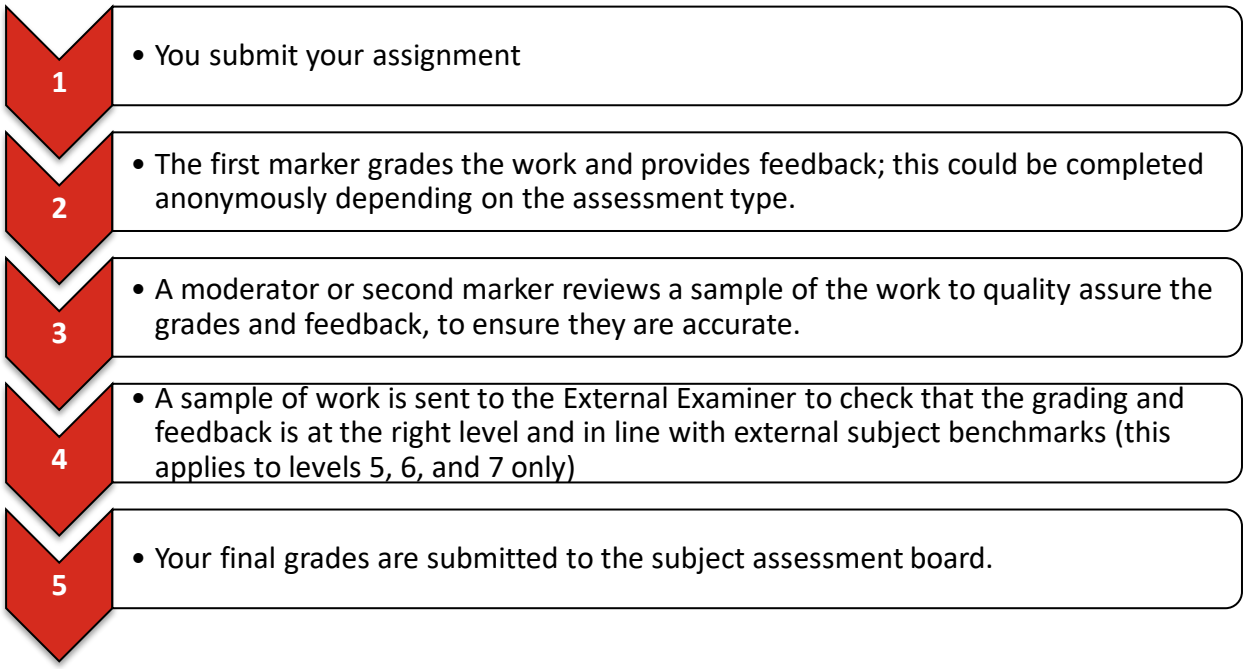
Further information is available at
https://unihub.mdx.ac.uk/study/assessment/regulations

Computer Science Mapping to the 20-point Scale					
Undergraduate	**Postgraduate**	**Grade**	**Percentages**		**Midpoint**
First	Distinction	1	80	100	90
		2	77	79	78
		3	73	76	74.5
		4	70	72	71
Upper second	Merit	5	68	69	68.5
		6	65	67	66
		7	63	64	63.5
		8	60	62	61
Lower second	Pass	9	58	59	58.5
		10	55	57	56
		11	53	54	53.5
		12	50	52	51
Third		13	48	49	48.5
		14	45	47	46
		15	43	44	43.5
		16	40	42	41
Compensatable fail	Compensatable fail	17	38	39	37
		18	35	37	36
Uncompensatable fail	Uncompensatable fail	19	0	34	17

Note that this system does not apply to modules which are graded pass / fail.

Assessment process

The following diagram provides an overview of the marking process for your module assessment. Details of the programme external examiner can be found in the programme handbook. Further information on the role of external examiners can be found at. http://unihub.mdx.ac.uk/your-study/ensuring-quality/external-examiners



Anonymous Marking Assessment Policy

We have worked with the MDXSU to create an anonymous marking policy, in response to student feedback. Anonymous marking ensures that your identity (your name, student number and other personal/identifiable information) is not made available to academics when they are marking your work. This means that you can have confidence that your assessments will be marked fairly and consistently. However, there are some forms of assessment for which anonymity cannot be guaranteed and these are recognised in the policy. We believe that it is important to provide you with the support and guidance needed to help you develop and prepare for your final assessments (those which count towards your final grades i.e. summative assessments). Therefore, anonymous marking will not apply to learning activities and assessments that do not contribute to your final grades (i.e. formative assessments). If you require further information and support to understand how anonymous marking works in your programme modules please contact the Module Leader for more information.

The Anonymous Marking Assessment Policy is available at:

https://www.mdx.ac.uk/__data/assets/pdf_file/0037/563599/anonymous-marking-assessment-policy.pdf