

This Version is No Longer Current
 The latest version of this module is available [here](#)

MODULE DESCRIPTOR

Module Title

Big Data Systems

Reference	CM3708	Version	1
Created	February 2019	SCQF Level	SCQF 9
Approved	May 2019	SCQF Points	30
Amended		ECTS Points	15

Aims of Module

To introduce students to the use of state-of-the-art Big Data analytics techniques and tools, including NoSQL data stores, and modern parallel computation methodologies.

Learning Outcomes for Module

On completion of this module, students are expected to be able to:

- 1 Compare and contrast different types of NoSQL data stores.
- 2 Critically analyse the suitability of a NoSQL data store for a given problem.
- 3 Extract actionable knowledge from big data, using the parallel computation framework.
- 4 Design, implement and evaluate scalable program solutions using a big data computation framework.

Indicative Module Content

NoSQL data stores (e.g., key-value, document, and graph). Case studies of NoSQL data stores. Properties of NoSQL data stores. Schema migration in NoSQL data stores. Modern parallel data processing techniques, e.g. MapReduce/Hadoop, Spark. Case studies on using parallel data processing for analysis and mining of Big Data.

Module Delivery

The module is delivered in Blended Learning mode using structured online learning materials/activities and directed study, facilitated by regular online tutor support. Workplace Mentor support and work-based learning activities will allow students to contextualise this learning to their own workplace. Face-to-face engagement occurs through annual induction sessions, employer work-site visits, and modular on-campus workshops.

Indicative Student Workload

	Full Time	Part Time
Contact Hours	30	N/A
Non-Contact Hours	30	N/A
Placement/Work-Based Learning Experience [Notional] Hours	240	N/A
TOTAL	300	N/A
<i>Actual Placement hours for professional, statutory or regulatory body</i>	240	

ASSESSMENT PLAN

If a major/minor model is used and box is ticked, % weightings below are indicative only.

Component 1

Type: Coursework Weighting: 50% Outcomes Assessed: 3, 4
 Description: This coursework will consist of a big data development exercise.

Component 2

Type: Practical Exam Weighting: 50% Outcomes Assessed: 1, 2
 Description: This practical exam will consist of a presentation on aspects of big data within the workplace.

MODULE PERFORMANCE DESCRIPTOR**Explanatory Text**

The calculation of the overall grade for this module is based on 50% weighting of C1 and 50% weighting of C2. An overall minimum grade of D is required to pass the module.

		Practical Exam:						
		A	B	C	D	E	F	NS
Coursework:	A	A	A	B	B	C	E	
	B	A	B	B	C	C	E	
	C	B	B	C	C	D	E	
	D	B	C	C	D	D	E	
	E	C	C	D	D	E	E	
	F	E	E	E	E	E	F	
	NS	Non-submission of work by published deadline or non-attendance for examination						

Module Requirements

Prerequisites for Module	None, in addition to course entry requirements.
Corequisites for module	None.
Precluded Modules	None.

INDICATIVE BIBLIOGRAPHY

- 1 HARRISON, G., 2015. Next Generation Databases: NoSQL, NewSQL, and Big Data. Apress.
- 2 LESKOVEC, J., ANAND, R. and ULLMAN, J.D., 2015. Mining of massive datasets. Cambridge University Press.
- 3 BERMAN, J., 2018. Principles and practice of big data: preparing, sharing, and analyzing complex information. 2nd ed. London: Academic Press.
- 4 MISHRA, R., 2018. PySpark recipes: a problem-solution approach with PySpark2. United States: Apress.
- 5 WIKTORSKI, T., 2019. Data-intensive systems principles and fundamentals using Hadoop and Spark. Cham: Springer.