Module Specification

Module Summary Information

1	**Module Title**	Modern Data Stores
2	**Module Credits**	20
3	**Module Level**	6
4	**Module Code**	CMP6207

5	**Module Overview**

This module will introduce NoSQL databases and distributed data storage frameworks through practice based activities. Through guided hands-on tutorials, you will become familiar with techniques using non-structured and semi-structured data examples. Also, you will gain knowledge on managing and processing data in a distributed infrastructure. This course will improve your development skills and provide experience with many different data systems

This module will be delivered by means of lectures, tutorials and practical lab work aimed at developing the knowledge and skills required to appreciate the issues, opportunities and practical approaches. While the lectures will cover the theory and concepts, the labs will cover the practical aspects. You are expected to investigate the topics before sessions. This approach will enable you to gain an insight into how modern data stores are being used in organisations before you learn supporting theory and practical application through the use of software. You are expected to come to sessions prepared and having completed all the exercises and activities set.

Finally, this module is very practical and you will be expected to complete all set practical exercises in the order which they are presented. This approach will enable you to build your knowledge, skills and practice the embedded transferable employability skills.

6	**Indicative Content**

This module covers two topics; NoSQL databases and distributed data storage. The first part introduces non-relational data storage options and covers the creation, storage, access, and management of data in different storage options. NoSQL databases are generally built to be distributed and partitioned across many servers. They're built to scale out for high availability and to be flexible enough to handle semi-structured and unstructured data. Hence, the second part of the module will cover distributed file systems, frameworks for distributed processing of large data sets across clusters of computers using simple programming models like map/reduce, and big data storage solutions.

7	Module Learning Outcomes On successful completion of the module, students will be able to:	
	1	Appraise and effectively communicate the different principal types, theories and technologies associated with NoSQL databases.
	2	Critically compare, analyse and evaluate traditional Relational databases and NoSQL databases.
	3	Design, implement and professionally report a NoSQL application for a real enterprise or a complex case study together with sound justification for the approach adopted.
	4	Demonstrate sufficient knowledge and skills on installing, configuring and managing distributed data management frameworks.

8	Module Assessment			
Learning Outcome				
	Coursework	**Exam**	**In-Person**
1, 2, 4	x		
3			**X**

9	Breakdown Learning and Teaching Activities	
Learning Activities		**Hours**
Scheduled Learning (SL) includes lectures, practical classes and workshops, peer group learning, Graduate+, as specified in timetable		Lecture/Tutorial Topics (not necessarily delivered in the sequence below and some topics might need more than one tutorial): 1. Introduction: NoSQL databases 2. Graph databases 3. XML databases 4. Key-value stores and document databases 5. Key-value stores and document databases 6. Column stores 7. Object databases 8. Extensible record stores 9. Distributed database systems 10. Data fragmentation 11. Replication and synchronisation 12. Consistency **Lecture/Core Content Delivery:** 12 hours Practical Topics (not necessarily delivered in the sequence below and some topics might need more than one lab session): 1. Basics of JavaScript and JSON 2. Introduction to MongoDB 3. Creating, Updating and Deleting Documents in MongoDB 4. Querying in MongoDB 5. Indexing in MongoDB 6. Aggregation in MongoDB 7. Replication and Sharing in MongoDB 8. Installing and Configuring Apache Hadoop 9. Running MapReduce in Hadoop

	Laboratory/Practical Demonstration/Workshop: 24 hours **Total:** 48 hours
Directed Learning (DL) includes placements, work-based learning, external visits, on-line activity, Graduate+, peer learning, as directed on VLE	Independent Study : 72 hours
Private Study (PS) includes preparation for exams	80 hours
Total Study Hours:	200 hours