

Module Specification

Module Summary Information

1	Module Title	Big Data Management
2	Module Credits	20
3	Module Level	7
4	Module Code	CMP7203

5	Module Overview
<p>This module focuses on aspects of managing big data systems with respect to the five V's (Volume, Velocity, Variety, Veracity, and Valence); i.e. systems that provide operational capabilities for real-time, interactive workloads where data is primarily captured and stored to support any analytical capabilities.</p> <p>Big Data has driven the creation of new technology architectures with the likes of NoSQL, MPP databases and Hadoop that enable new types of products and services. Operational systems, such as the NoSQL databases, focus on servicing highly concurrent requests while exhibiting low latency for responses operating on highly selective access criteria.</p> <p>Learning and Teaching</p> <p>This module will introduce big data management techniques through practice based activities associated with real world issues. You will gain knowledge pertaining to the application of big data management techniques through practical experience, theory in practice and by exploring 'real world' issues via a range of software proprietary and open source big data analytics tools.</p> <p>The module is designed as a series of pre-reading and discussions, interactive-taught lectures with breakout sessions and workshop/lab exercises using a variety of tools and techniques. You are expected to investigate the topics before sessions. This approach will enable you to gain an insight into how specific big data management techniques are being used in organisations before you learn supporting theory and practical application through the use of big data management software. You are expected to come to sessions prepared and having completed all the exercises and activities set.</p> <p>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.</p>	

6	Indicative Content
<p>Introduction to Data Warehousing Architecture Dimension modelling ETL Process OLAP and SQL Business Intelligence</p> <p>Introduction to Big Data: Beginning and examples of Big Data; value in Big Data (Example Applications); classify sources of big data; structured vs unstructured data; 5 Vs of Big Data</p>	

Introduction to Data Models: Relational vs semistructured Data Model, key value, graph Data Model, column-family

Introduction to Big Data Management Systems: Distributed File Systems,; Hadoop Ecosystem; Mapreduce; Oracle Cloud Computing as enabler; Understanding Hadoop commands; Running sample code on Hadoop; Oracle Big Data Connectors between database and Hadoop Clusters

Storage and Management
 DBMS vs non-DBMS storage
 Data governance and ethics

SQL for Big Data (Hadoop) and Introduction to NoSQL
 Oracle NoSQL / Open Source equivalent

7	Module Learning Outcomes		
	On successful completion of the module, students will be able to:		
	1	Critically evaluate modern big data processing paradigms	
	2	Develop and implement a big data solution.	
	3	Analyse use cases, visualise and report the results of a big data solution	
	4	Assess how ethics govern the design choices in devising a Big Data enabled solution.	

8	Module Assessment		
Learning Outcome			
	Coursework	Exam	In-Person
1-4	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	48	
Directed Learning (DL) includes placements, work-based learning, external visits, on-line activity, Graduate+, peer learning, as directed on VLE	76	
Private Study (PS) includes preparation for exams	76	
Total Study Hours:	200	