

Dublin Business School Module Descriptor							
<b>Stage</b>		Award			<b>Semester</b>		L
<b>Module Title</b>		Data Management & Analytics					
<b>Module Status</b>		Elective					
<b>Module Code</b>		<b>Pre-Requisite Module Code (s)</b>		<b>Co-Requisite Module Code(s)</b>		<b>Level</b>	<b>Credits</b>
M3.12		M2.25				8	10
		<b>Date Approved</b>		<b>Date for Review</b>			<b>Capstone</b>
<b>Contact Hours</b>				<b>Non-contact Hours</b>			<b>Total Effort</b>
<b>Lecture</b>	<b>Tutorial</b>	<b>Practical</b>	<b>Seminar</b>	<b>Assignment</b>	<b>Placement</b>	<b>Independent Work</b>	
36		36	3	75		100	250
Allocation of Marks Within the Module							
<b>Continuous Assessment</b>			<b>Project</b>	<b>Practical</b>	<b>Final Examination</b>		<b>Total</b>
50%					50%		100%

School of Business
<b>Author:</b> Michael Gleeson
<b>Description:</b>
High quality data and discovering patterns is an essential component of a modern organisation. Enterprises now rely on their data assets to make informed and increasingly effective decisions. Data analytics involves processes and activities designed to obtain and extract useful information from available data stores. The emergence of large scale data sets (referred to as 'Big Data') enforced a reassessment of traditional data management and analytic techniques. This module will examine core data management tools and techniques. It will then build on this knowledge to evaluate cutting edge tools and techniques.
<b>Aims:</b>
1. To give learners the ability to investigate data storage techniques and the process of information retrieval.
2. To develop in learners knowledge in the area of data warehousing and data mining techniques.
3. To allow learners to demonstrate data analysis techniques in order to discover patterns.
4. To give learners the knowledge to explain and implement emerging technologies in the area of Big Data.
5. To support learners to gain the experience to synthesise the qualities of Knowledge Management and Business Intelligence for an organisation.
6. To expose learners to real life application of theory through the use of case studies.

Learning Outcomes:
On successful completion of this module, learners will be able to:
1. Evaluate data storage systems and associated techniques for indexing and retrieving data.
2. Determine the best approach to designing a data warehouse and mining a variety of data sources.
3. Discover patterns and trends from data stores through the use of data analysis techniques.
4. Evaluate and implement modern large scale data processing techniques such as MapReduce
5. Evaluate the role and benefits of effective Knowledge Management and Business Intelligence in an organisation.

<b>Assessment Strategy:</b>
<b>Participant learning will be assessed by the following:</b>
<p>Learning will be assessed by a mixture of Formative and Summative Assessment as follows:</p> <ol style="list-style-type: none"> <li>1. Continuous assessment in the form of theory and lab based online assessment to allow the learner to gradually build knowledge throughout the module and encourage the learner to learn progressively.</li> <li>2. Continuous assessment in the form of a Skills Based Assessment (SBA) involving the development of an analytics solution for a particular data-set, used to establish and evaluate the learners' practical skills and abilities.</li> <li>3. A final end of year summative assessment to encompass all the modules theory and lab work.</li> </ol>

Method of Assessment	Percentage Weightings	Learning outcomes assessed
Online Assessment	30%	1, 2, 3
Skills Based Assessment	20%	4
Final Exam	50%	1,2,3,5

<b>Indicative Syllabus:</b>
<b>1.</b> Information as a Strategic Resource: The nature and use of information. Data as a business asset and a strategic enabler, data organisation issues. Problem of data fragmentation and isolation. Evolution of information usage and modern data models.
<b>2.</b> Data Storage Technologies: Data and information on a local system, files, file types, file systems and databases. Storage devices and storage configurations (clusters, SAN, NAS). Indexing features, XML basics and XML querying.
<b>3.</b> Web Centric Data: Information retrieval methodology and tools, indexing and searching techniques (bots, web crawlers, the deep web). Web 3.0 and semantic web approach, semantic web components (RDF, SPARQL and OWL).
<b>4.</b> Data Abstraction Layers: External, conceptual and internal schemata of data. Data independence, data definition and data manipulation languages. Relational data models, functional dependency theory and normalization.
<b>5.</b> Data Warehousing: Data warehouse features, OLTP v OLAP, data marts versus data warehousing, generic data warehouse architecture. Data warehouse applications, logical data warehouse design and data warehouse quality.
<b>6.</b> Data Mining: Discovering new knowledge, mining algorithms for modelling, classification, clustering, prediction, sequence analysis and association analysis. Transforming and preparing data for purposes of mining.
<b>7.</b> Data Analytics: Exploring data, basic concepts, decision trees, association analysis, cluster analysis and anomaly detection. Types and evaluation of analytic approaches. Social media, data quality and volume issues.
<b>8.</b> Data Intensive Computing: 'Big Data' era and scalable approaches to processing large amounts of information. HPC, parallel, cluster and grid computing, functional programming models including Hadoop technologies and Map/Reduce.
<b>9.</b> Large Scale Data Management: Context of large scale of data management, RDBMS, ACID properties and the CAP Theorem. Limitations of traditional methods, evolution of NoSQL non-relation data stores. BASE and types of NoSQL data stores.
<b>10.</b> Knowledge Management & Business Intelligence: Nature of knowledge management, knowledge capturing and modelling, Artificial Intelligence, Natural Language Processing. Business performance management, KPI's, dashboards, emerging trends in BI.
<b>11.</b> Data Security and Governance: Data security, the need for data security, data security policies, standards & procedures. Data usage, areas of risk, data policies, standards and regulatory

compliance.

**12.** Case Studies in Data Analytics: Examination and evaluation of practical application of data management techniques, including Web semantics, GIS and Multimedia mining using case studies.

<b>Indicative Bibliography</b>				
<b>Title</b>	<b>Author</b>	<b>Publisher</b>	<b>Published</b>	<b>Edition</b>
Introduction to Data Mining	Pang-Ning Tan	Prentice Hall	2013 (due to be published June)	2 <sup>nd</sup>
Big Data, Big Analytics	Michael Minelli	Wiley (CIO)	2013	1 <sup>st</sup>
Pro Hadoop	Jason Venner	Apress	2009	1 <sup>st</sup>
<b>Electronic Resources</b>				
DBS E-Learning Support	<a href="http://elearning.dbs.ie">http://elearning.dbs.ie</a>			
International Institute for Analytics	<a href="http://www.iianalytics.com/">http://www.iianalytics.com/</a>			
Data Mining & Analytics Resources	<a href="http://www.kdnuggets.com/">http://www.kdnuggets.com/</a>			
SAS Case Studies	<a href="http://www.sas.com/success/index.html">http://www.sas.com/success/index.html</a>			
DBS Library Resources	<a href="http://library.dbs.ie/Electronic-Resources/E-Library.htm">http://library.dbs.ie/Electronic-Resources/E-Library.htm</a>			