Module 24: Managing Big Data

Stage				4			
Semester				1			
Module Title				Managing Big Data			
Module Number				24			
Module Status				Elective			
Module ECTS Credits				5			
Module NFQ level				8			
Pre-Requisite Module Titles				Relational Databases			
Co-Requisite Module Titles				None			
Capstone Module?				No			
List of Module Teaching Personnel				Mr Mark Scanlon, Dr Waseem Akhtar			
Contact Hours				**Non-contact Hours**			**Total Effort (hours)**
36				**64**			**100**
Lecture	**Practical**	**Tutorial**	**Seminar**	**Assignment**	**Placement**	**Independent Work**	
12	24			24		40	
Allocation of Marks (Within the Module)							
	Continuous Assessment	**Project**	**Practical**	**Final Examination**			**Total**
Percentage Contribution	**40%**			**60%**			**100%**

Intended Module Learning Outcomes

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

1. Understand and perform the duties of a DBA
2. Implement query optimisation strategies
3. Recognise the important role of efficient transaction management with regards to concurrency control, database recovery and deadlock detection/prevention.
4. Understand the considerations surrounding the processing of Big Data
5. Describe and implement various strategies in Data Mining and Warehousing
6. Describe the use of Data Analytics on Big Data

Module Objectives

There are two aims to this module: to expose you to practical issues in database management systems such as database administration and query optimisation; and to give you a flavour of the procedures and considerations in handling Big Data. In order to gain an understanding of how to work with Big Data, you gain an understanding of the core concepts required such as Data Mining, Data Warehousing and Data Analytics.

Module Curriculum

Database Management
The role of the DBA / Security / User Management / Physical Database Issues

Query Optimisation
Use of indexing and keys / Optimising Joins / Optimising queries in a RDBMS

Transaction Processing and Concurrency
Transactions Stages; commit, abort, etc. / ACID properties / Concurrency problems; lost update, incorrect summary, dirty read etc. / Locking / Deadlock detection and prevention

Introduction to Big Data
Data Model / Data Storage / Data Warehousing / Data Extraction, Transforming and Loading / Batch Processing / Scalability / NoSQL / Managing Big Data, Online Analytical Processing

Data Mining
Structural Pattern Recognition / Input-Output / Clustering / Managing Data Warehousing Models (Bottom-Up, Top-Down, etc.) / Data Transformation Models

Introduction to Data Analytics
Extracting information from Big Data / Statistics / Case Studies

Reading Lists and other learning materials

Recommended Reading

Big Data - Principles and best practices of scalable realtime data systems	Marz, N. and Warren, J.	Manning Publications	2013
Fundamentals of Database Systems (6th Edition)	Elmasri, R. and Navathe, S. B.	Addison-Wesley	2010
Data Mining: Practical Machine Learning Tools and Techniques (3rd Edition)	Witten, I., Frank, E. and Hall, M.	Morgan Kaufmann Publishers	2011

Additional reading as recommended by lecturer, appropriate to topic.

Module Learning Environment

Lectures are carried out in classrooms / lecture halls in the College. Computer Labs throughout the Campus are accessible for the purpose of completing assignments.

Library

All learners have access to an extensive range of physical and electronic (remotely accessible) library resources. The library monitors and updates its resources on an on-going basis, in line with the College's Library Acquisition Policy. Lecturers update reading lists for this course on an annual basis, as is the norm with all courses run by Griffith College.

Module Teaching and Learning Strategy

Classes are used to explain the concepts, exemplify the techniques, and solve (in workshop style) a series of exercises and problems.

In addition to classes, you need to put in at least four hours of study and homework each week.

Module Assessment Strategy

Name	Description	Weighting	Learning Outcomes
Assignment 1	Query optimisation assignment	20	1, 2
Assignment 2	Big Data/Data Warehousing assignment	20	4, 5
Examination	End of term exam	60	All