

COURSE INFORMATION

School/Faculty:	Computing/Engineering	Page:	1 of 5
Program name:	Master of Science (Data Science)		
Course code:	MCDS 1123	Academic Session/Semester:	2023/24/1
Course name:	Big Data Management	Pre/co requisite (course name and code, if applicable):	
Credit hours:	3		

Course synopsis	This course provides a basic fundamental of big data architecture and management. Students will learn the big data processes and the current big data technologies that are available. Further, students will be exposed to the big data platform ecosystem for big data manipulation. The big data management will be explored for the best practice in managing and manipulating large amount of data. At the end of the course, students should be able to understand the architecture and management of big data and also can develop simple application of big data handling using particular platform in assignment.			
Course coordinator (if applicable)	Assoc Prof Dr Mohd Shahizan Othman (shahizan@utm.my)			
Course lecturer(s)	Name	Office	Contact no.	E-mail

Mapping of the Course Learning Outcomes (CLO) to the Programme Learning Outcomes (PLO), Teaching & Learning (T&L) methods and Assessment methods:

No.	CLO	PLO (Code)	*Taxonomies and **generic skills	T&L methods	***Assessment methods
CLO1	Understand the technology for managing, processing and manipulating large amount of data.	PLO1	C3, A3	Lecture, active learning, lab	Asg, F
CLO2	Design big data platform demonstrating the implementation of big data applications.	PLO3	C6	Lecture, active learning, lab	Asg, F, PR
CLO3	Discuss current technology that support for sustainability of the big data platform ecosystem.	PL06	C3, LL2	Lecture, Active Learning, Project based learning	Asg, PR, F

Prepared by:	Certified by:
Name:	Name:
Signature:	Signature:
Date:	Date:

School/Faculty:	Computing/ Engineering	Page:	2 of 5
Program name:	Master of Science (Data Science)		
Course code:	MCDS 1123	Academic Session/Semester:	2023/24/1
Course name:	Big Data Management	Pre/co requisite (course name and code, if applicable):	
Credit hours:	3		

Refer *Taxonomies of Learning and **UTM's Graduate Attributes, where applicable for measurement of outcomes achievement ***T – Test; Q – Quiz; HW – Homework; Asg – Assignment; PR – Project; Pr – Presentation; F – Final Exam etc.					

Details on Innovative T&L practices:

No.	Type	Implementation
1.	Active learning	Conducted through in class activities such case study discussion and hands on technology
2.	Project-based learning	Conducted through group project. Students in a group of 3 are required to analyse and design an appropriate big data management tool.

Weekly Schedule:

Week 1	Introduction to Big Data and Big Data Analytic Introduction to big data fundamentals and concepts
Week 2	Big Data Processing and Technology Batch, Real time and streaming processing Big Data Processing issues (Scalability, Storage, Sourcing)
Week 3 -4 Assignment 1	Big Data Processing and Technology ACID, BASE and CAP theorem Distributed File Processing & Map Reduce Processing
Week 5	Big Data Processing and Technology Lambda Architecture
Semester Break 21-29 April 2023	
Week 6-7 Assignment 2	Relational Database (RDBMS) Relational Data Modelling Database design phases
Semester Break 4-10 Dec 2022	
Week 8 Lab assignment 1	Relational Database (RDBMS) SQL programming (DDL, DML, CRUD Operation)
Week 9 Lab assignment 2	Relational Database (RDBMS) SQL programming (Subqueries, Join Tables, Aggregate)

School/Faculty:	Computing/ Engineering	Page:	3 of 5
Program name:	Master of Science (Data Science)		
Course code:	MCDS 1123	Academic Session/Semester:	2023/24/1
Course name:	Big Data Management	Pre/co requisite (course name and code, if applicable):	
Credit hours:	3		

Week 10	No SQL database Introduction to No SQL database Semi structured data Modelling (Key Value, Column Family, Document and Graph)
Week 11-12 Lab assignment 3 & 4	No SQL database (Document-based Database) Document-based data modelling MongoDB query language
Week 13 - 14	Cloud Technology Introduction to Cloud AWS cloud via AWS Learning Management System)
Week 15	Project Presentation

Transferable skills (generic skills learned in course of study which can be useful and utilised in other settings):

Live long learning

Student learning time (SLT) details:

Distribution of student Learning Time (SLT) Course content outline					Teaching and Learning Activities			TOTAL SLT
	Guided Learning (Face to Face)				Guided Learning Non-Face to Face	Independent Learning Non-Face to face		
CLO	L	T	P	O				
CLO1	9h		3h	1h	10h	10h		33h
CLO2	10h			2h	12h	12h		36h
CLO3	11h		3h	3h	14h	14h		45
Total SLT	30h		6h	6h	36h	36h		114

Continuous Assessment		PLO	Percentage	Total SLT
1	Assignment 1	PLO1	10	As in CLO1
2	Assignment 2	PLO1	10	As in CLO1 – (5 h)
4	Lab Assignment		15	
5	AWS learning		5	
5	Project 1	PLO3, PLO6	15	As in CLO2, CLO3-(10h)
6	Project 2		15	

School/Faculty:	Computing/ Engineering	Page:	4 of 5
Program name:	Master of Science (Data Science)		
Course code:	MCDS 1123	Academic Session/Semester:	2023/24/1
Course name:	Big Data Management	Pre/co requisite (course name and code, if applicable):	
Credit hours:	3		

Final Assessment			Percentage	Total SLT
1	Final Exam	PLO3	30	3h
Grand Total			100	120h

L: Lecture, T: Tutorial, P: Practical, O: Others

Special requirement to deliver the course:

- 1) Computer Lab for Lab Session
- 2) SQL software, No SQL database

Learning resources:

Text Book (If applicable)

-

Main references

1. Nathan Marz and James Warren, Big Data: Principle and Best practises of scalable real time data system, Manning Publication, 2015
2. Alan Beaulieu, Learning SQL: Generate, Manipulate and Retrieve Data from Relational Database, Third Edition, O'Reilly Media, 2020.
3. Ray Rafaels, Cloud Computing: From Beginning to End, MIT Press 2018
4. Martin Kleppmann, Designing Data-Intensive Application: The Big Ideas Behind Reliable, Scalable and Maintaining Systems, O'Reilly Media 2017
5. Shannon Bradshaw and Kristina Chodorow, MongoDB: The Definite Guide 3rd Edition, O'Reilly Media 2020
6. W.H Inmon, Daniel Linstedt and Mary Levins, Data Architecture: A Primer for the Data Scientist, 2nd edition Elsevier / Morgan Kaufmann, 2019.

Additional Reference

Online

<http://elearning.utm.my>

<https://azure.microsoft.com/en-us/overview/what-is-azure/>

Academic honesty and plagiarism: (Below is just a sample)

Copying of work from other students/groups or from other sources is not allowed. Brief quotations are allowed and then only if indicated as such. Existing texts should be reformulated with your own words used to explain what you have read. It is not acceptable to retype existing texts and just acknowledge the source as a reference. Be warned: students who submit copied work will obtain a mark of **zero** for the assignment and disciplinary steps may be taken by the Faculty. It is also unacceptable to do somebody else's work, to lend your work to them or to make your work available to them to copy.

Other additional information (Course policy, any specific instruction etc.):

School/Faculty:	Computing/ Engineering	Page:	5 of 5
Program name:	Master of Science (Data Science)		
Course code:	MCDS 1123	Academic Session/Semester:	2023/24/1
Course name:	Big Data Management	Pre/co requisite (course name and code, if applicable):	
Credit hours:	3		

-

Disclaimer:

All teaching and learning materials associated with this course are for personal use only. The materials are intended for educational purposes only. Reproduction of the materials in any form for any purposes other than what it is intended for is prohibited.

While every effort has been made to ensure the accuracy of the information supplied herein, Universiti Teknologi Malaysia cannot be held responsible for any errors or omissions.