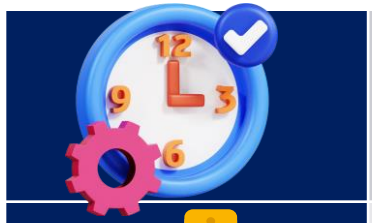
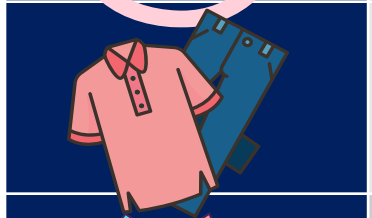
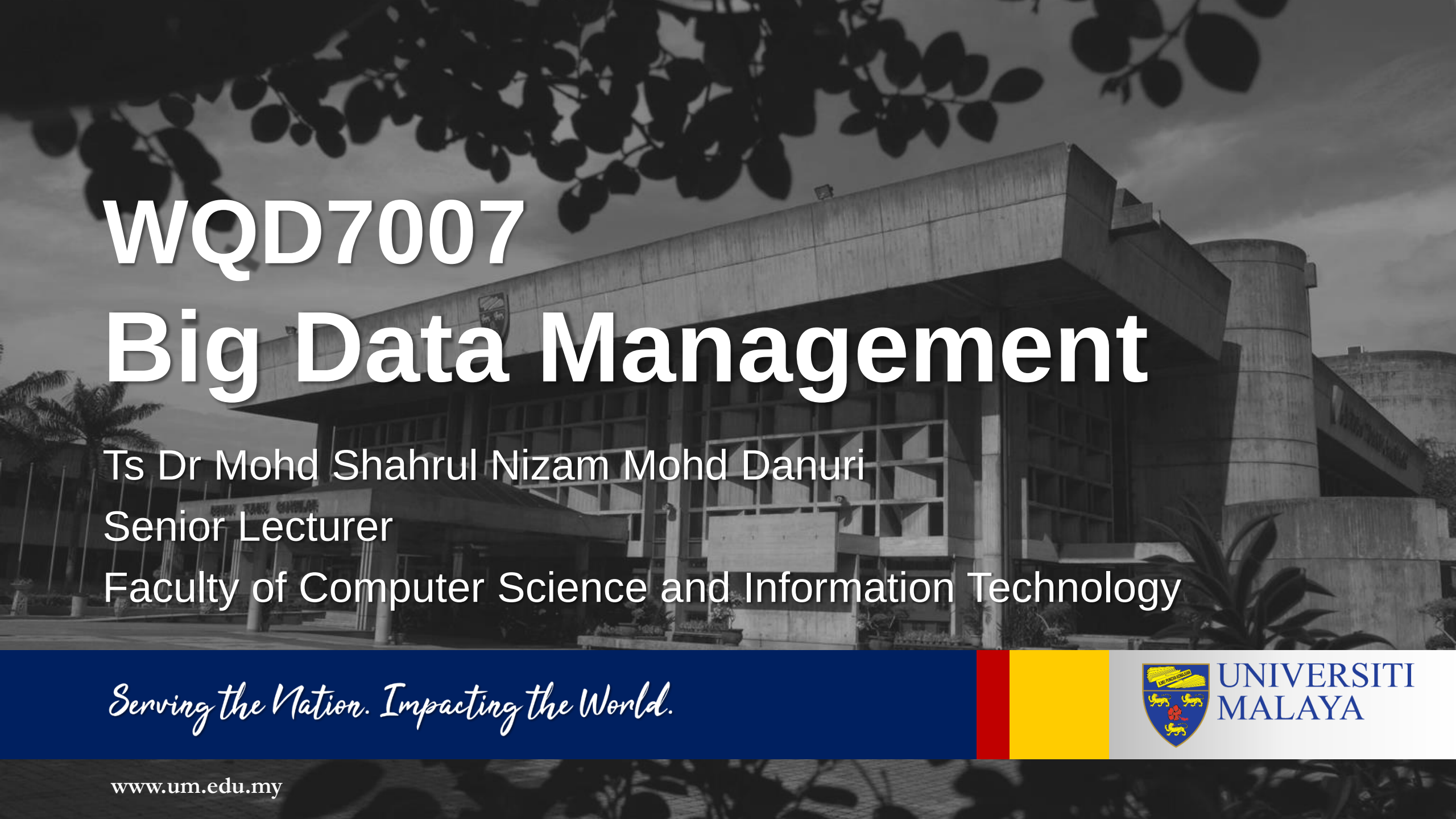
# CLASSROOM REMINDERS

| | |
|---|---|
|  | **BE ON TIME**<br>Set up early so that you're good and ready when class starts. |
|  | **BE PREPARED FOR CLASS**<br>Have everything you need within reach. |
|  | **OPEN YOUR CAMERA, OFF YOUR MIC**<br>Make sure you look presentable for class. |
|  | **WEAR APPROPRIATE CLOTHES**<br>Make sure you look presentable for class. |
|  | **BE RESPECTFUL AND ONLINE ETIQUETTE**<br>Respect your teacher and classmates. Good behavior and manners expected of us when we're in class |



# WQD7007
# Big Data Management

Ts Dr Mohd Shahrul Nizam Mohd Danuri
Senior Lecturer
Faculty of Computer Science and Information Technology

# OTHER CLASSROOM REMINDERS

| Don't be afraid to ask questions | Be courteous and polite | Keep up with lessons and homework |
|---|---|---|
| Feel free to ask your teacher if you don't know something or if something is unclear. | Think before speaking. Say "thank you" and "please" when appropriate. | Review your notes. Stick to a study routine so you won't get behind with class assignments. |
| **Submit assignments correctly** | **Take breaks from being online** | **Support each other** |
| Follow your teacher's rules and instructions on how to submit your work online. | Find time to recharge and engage in offline activities. Find balance and make time for your hobbies as well. | Be kind to everyone. Let's encourage and lift each other up. |

# Who Am I?

- Ts. Dr. Mohd Shahrul Nizam bin Mohd Danuri
- 14 years in ICT industry | 10 years in Education
  - Software development, networking, telecommunication, research and technology in agriculture
  - » PhD in Information Management, UiTM
  - » Master's Degree in Intellectual Property, UKM
  - » BSc. Computer Science (Hons), USM
  - » MBOT, IEEE, MNCC, Mendeley Advisor
  - » Program Usahawan Teknikal MARA, Program Teknousahawan MARA-SIRIM
- Room: B-3-18 [Block B, FCSIT]
- Mobile: +6012.636.6441 | +603.7967.6399
- Email: msnizam@um.edu.my

# Who Am I?

**Network**

tmnet Let's click

streamyx ZONE

TM WiFi

**Programmer**

```
<CIF code>
        <Display "Hello World">
</CIF code>
```

** All images belong to respective owners*

**Computer Engineering**

# Who Am I?

Information Systems

Big Data / Data Science

Internet of Things (IoT)

Wireless Sensor Network (WSN)

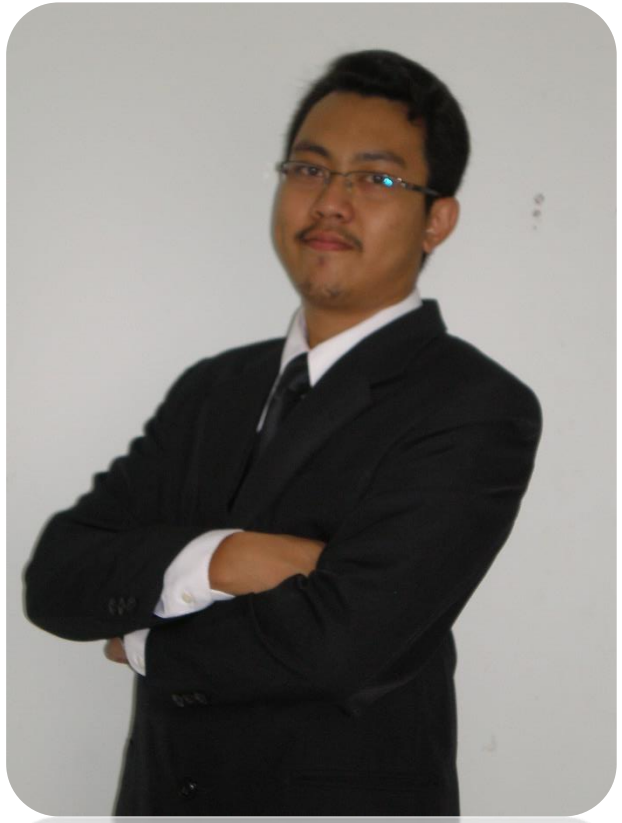
Technology in Agriculture

https://www.msnizam.net

# Who Am I?

- Senior Lecturer
- Faculty of Computer Science and Information System
- Universiti Malaya

- Internship Coordinator - IS
- Project Manager - FSKTM Technovation



# My Career Path (2001-2023)

Lecturer & Researcher

Chief Technology Officer / Chief Executive Officer

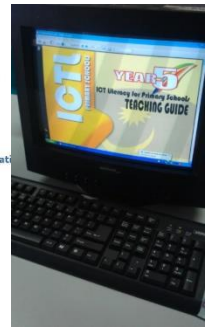
(Employer)



Support Engineer/ Pre-sales Consultant (Manager)

System Engineer (Software/ Network/ etc.)

System Engineer (Programme



https://itskkulai.wordpress.com/2012/07/17/modul-literasi-ict-ictl-tahun-5-2012/

# Course Information

# Academic Calendar

| SEMESTER II | | | | | |
|---|---|---|---|---|---|
| Lectures | 7 | weeks* | 17.03.2025 | - | 04.05.2025 |
| Mid Semester II Break | 1 | week | 05.05.2025 | - | 11.05.2025 |
| Lectures | 7 | weeks* | 12.05.2025 | - | 29.06.2025 |
| Revision Week | 1 | week* | 30.06.2025 | - | 06.07.2025 |
| Semester II Final Examination | 3 | weeks* | 07.07.2025 | - | 27.07.2025 |
| Semester II Break | 4 | weeks | 28.07.2025 | - | 24.08.2025 |
| | 23 | weeks | | | |
| **SPECIAL SEMESTER** | | | | | |
| Lectures | 7 | weeks* | 28.07.2025 | - | 14.09.2025 |
| Special Semester Final Examination | 1 | week* | 15.09.2025 | - | 21.09.2025 |
| Break | 1 | week | 22.09.2025 | | 28.09.2025 |
| | 9 | weeks | | | |

***Note:***

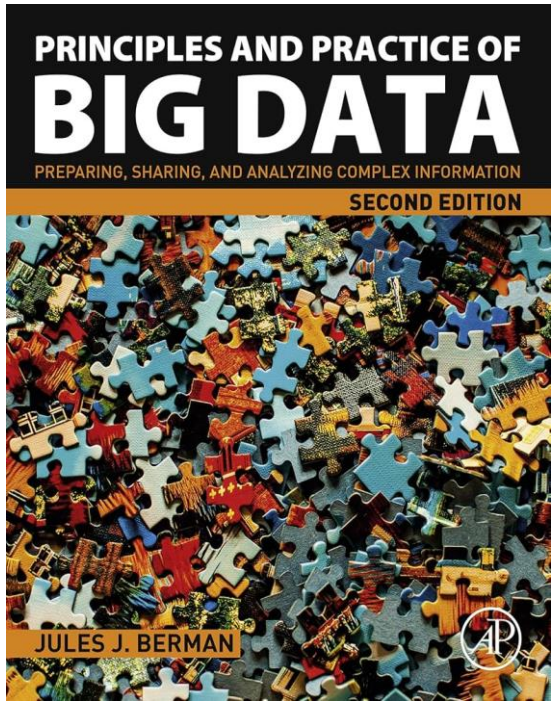
*(*) The Academic Calendar has taken into account public and festive holidays and is subject to change:*

| | |
|---|---|
| Deepavali | 01 November 2024 (Friday) |
| Christmas Day | 25 December 2024 (Wednesday) |
| New Year | 01 January 2025 (Wednesday) |
| Chinese New Year | 29 & 30 January 2025 (Wednesday & Thursday) |
| Federal Territory Day | 01 February 2025 ((Saturday) |
| Thaipusam | 11 February 2025 (Tuesday) |
| Nuzul Al-Quran | 17 March 2025 (Monday) |
| Eidul Fitri | 31 March & 01 April 2025 (Monday & Tuesday) |
| Wesak Day | 12 May 2025 (Monday) |
| His Majesty the King's Birthday | 02 June 2025 (Monday) |
| Eidul Adha | 06 June 2025 (Friday) |
| Awal Muharam | 27 June 2025 (Friday) |

Senate Approval Date: 24.01.2024

# Textbook

- Main textbook
- Most of the resources is available in the book (case study, etc)
- Relying only on lecture slides provided is NOT enough!




# Course Learning Outcome

- CLO1: Explain the processes in data pipeline.
- CLO2: Discuss database concepts and technologies for big data storage and retrieval.
- CLO3: Apply appropriate models, tools, and technologies to implement storage, search and retrieval systems for large-scale structured and unstructured information systems.
- CLO4: Analyse data provenance and data trustworthiness, and its role in sharing and reuse of data.

# Course Requirements

- Penilaian Berterusan / Continuous Assessment : 60%
  - » Tugasan (Minggu 4-5) / Assignment (Week 4-5) : **15%**
  - » Ujian pertengahan (Minggu 8) / Mid-term Test (Week 8) : **10%**
  - » Ujian Makmal (Minggu 11) / Lab test (Week 11) : **15%**
  - » Projek (Minggu 13-14) / Project (Week 13-14) : **20%**
- Penilaian Akhir / Final Assessment: 40%
  - » Penilaian Alternatif 1 (Minggu 14) / Alternative assessment 1 (week 14) : **15%**
  - » Penilaian Alternatif 2 (Minggu 14) / Alternative assessment 2 (week 14) : **25%**

# Topics Covered

- Big data pipeline using Hadoop
- Big data concepts
- Big data technologies
- Distributed computing for big data
- Data provenance and data trustworthiness

# Teaching Plan

| | |
|---|---|
| Sesi Akademik<br>*Academic Session* | 2024/2025 |
| Semester/Penggal<br>*Semester/Term* | 2 |
| Kod Kursus<br>*Course Code* | WQD7007 |
| Tajuk Kursus<br>*Course Title* | Pengurusan Data Besar<br>*Big Data Management* |
| Bahasa Pengantar<br>*Medium of Instruction* | Bahasa Inggeris<br>*English* |
| Rujukan Utama<br>*Main Reference* | 1. Principles and Practice of Big Data: Preparing, Sharing, and Analyzing Complex Information, Jules J. Berman, 2018.<br>2. Big Data Systems: A 360-degree Approach, Jawwad Ahmed Shamsi and Muhammad Ali Khojaye, 2021.<br>3. Big Data Fundamentals: Concepts, Drivers, & Technique, Wajid Khattak and Paul Buhler, 2016.<br>4. MongoDB: The Definitive Guide, Powerful and Scalable Data Storage. Shannon Bradshaw and Eoin Brazil et al., 2019.<br>5. Big Data Governance: Modern Data Management Principles for Hadoop, NoSQL & Big Data Analytics, Peter Ghavami, 2016.<br>6. Modern Big Data Processing with Hadoop, V Naresh Kumar & Prashant Shindgikar, 2018. |
| Strategi Pembelajaran<br>*Learning Strategies* | Segerak (F2F) Kuliah, tutorial dan makmal<br>Segerak (NF2F) forum atas talian<br>Pelajaran tidak segerak: Kuliah pra-rakaman, tutorial pra-rakaman, makmal pra-rakaman<br><br>*Synchronous (F2F): Lecture, tutorial and lab*<br>*Synchronous (NF2F): Online forum*<br>*Asynchronous lessons: Pre-recorded lecture, pre-recorded tutorial, pre-recorded lab* |
| Masa Pembelajaran Pelajar<br>*Student Learning Time* | Bersemuka / *Face to face* : 24<br>Tidak Bersemuka / *Non Face to face*: 18<br>Masa Persediaan Pelajar / *Student Preparation Time*: 78 |
| Kemahiran Boleh Pindah<br>Transferable Skills | Penyelesaian masalah dengan menggunakan kaedah - kaedah statistik, kemahiran mengintepretasi keputusan, dan kemahiran analisis/logikal;<br><br>*Problem solving using stastistical technique, result evaluation skills, and analysis/logical skills.* |
| Pensyarah / *Lecturer*<br>Bilik / *Room*<br>Telefon/e-mel<br>*Telephone/e-mail* | OCC3: Ts. Dr. Mohd Shahrul Nizam Mohd Danuri, B-3-18, 03-79676399, msnizam@um.edu.my (Course Coordinator)<br><br>OCC1 & 2: Assoc. Prof. Dr. Tutut Herawan, B-3-5, 03-99672509, tutut@um.edu.my |
| Sesi Kuliah / *Lecture Session:*<br><br>Hari/Masa / *Day/Time*<br>Tempat / *Venue* | OCC1: Selasa, 6pm-8pm / *Tuesday, 6pm-8pm*<br>OCC2: Sabtu, 8am-10am / *Saturday, 8am-10am*<br>OCC3: Sabtu, 3pm-5pm / *Saturday, 3pm-5pm*<br><br>Online (MS Team link will be provided in Spectrum) |
| Sesi Tutorial/Amali:<br>*Tutorial/Practical Session:*<br><br>Hari/Masa / *Day/Time*<br>Tempat / *Venue* | OCC1: Selasa, 8pm-9pm / *Tuesday, 8pm-9pm*<br>OCC2: Sabtu, 10am-11am / *Saturday, 10am-11am*<br>OCC3: Sabtu, 5pm-6pm / *Saturday, 5pm-6pm*<br><br>Online (MS Team link will be provided in Spectrum) |
| Perincian Pemberatan Penilaian<br>*Detail of Assessment Weightage* | Penilaian Berterusan / *Continuous Assessment* : 60%<br>- Tugasan (Minggu 4-5) / *Assignment (Week 4-5)* : 15%<br>- Ujian pertengahan (Minggu 8) / *Mid-term Test (Week 8)* : 10%<br>- Ujian Makmal (Minggu 11) / *Lab test (Week 11)* : 15%<br>- Projek (Minggu 13-14) / *Project (Week 13-14)* : 20%<br>Peperiksaan Akhir / *Final Examination* : 40%<br>- Penilaian Alternatif 1 (Minggu 14) *Alternative assessment 1 (week 14)* : 15%<br>- Penilaian Alternatif 2 (Minggu 14) *Alternative assessment 2 (week 14)* : 25% |

# Teaching Plan

**Jadual Pengajaran / *Teaching Schedule***

| Minggu<br>*Week* | Topik & Aktiviti<br>*Topic & Activities* | Rujukan<br>*References* |
|---|---|---|
| 1 | Topik/ *Topic*:<br>Pengenalan kepada kursus:<br>Data Besar dari kacamata Dunia Perniagaan & Dunia Sains<br>*Introduction to the course:*<br>*Big Data from Business and Science perspective*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Tutorial (1 jam) / *tutorial (1 hour)* | Bahan atas talian<br>*Online material* |
| 2 | Topik/ *Topic*:<br>Talian paip Data Besar dengan *Hadoop*<br>*Big Data Pipeline using Hadoop*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Makmal (1 jam) / *Lab (1 hour)* | Bahan atas talian<br>*Online material* |
| 3 | Topik/ *Topic*:<br>Konsep-konsep pangkalan data bagi data besar: Menyediakan struktur kepada data tidak berstruktur<br>*Big data concepts: providing structure for unstructured data*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Tutorial (1 jam) / *tutorial (1 hour)* | Berman (2013)<br><br>Bahan atas talian<br>*Online material* |
| 4 | Topik/ *Topic*:<br>Konsep-konsep pangkalan data bagi data besar: Pengenalpastian ontologi<br>*Big data concepts: Identification and ontologies*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Makmal (1 jam) / *Lab (1 hour)*<br>Penghantaran tugasan (PDF melalui Spectrum, 10 muka surat, 15%) / *Assignment submission (PDF through Spectrum, 10 pages, 15%)* | Berman (2013)<br><br>Bahan atas talian<br>*Online material* |
| 5 | Topik/ *Topic*:<br>Teknologi penyimpanan Data Besar: Capaian data tradisional & pengurusan pangkalan data tradisional<br>*Big data technologies: Traditional database management*<br>Aktiviti / *Activities*:<br>NF2F: Kuliah pra-rakaman (2 jam) / *Pre-recorded Lecture (2 hours)*<br>NF2F: Tutorial pra-rakaman (1 jam) / *Pre-recorded tutorial (1 hour)* | Berman (2013)<br><br>Bahan atas talian<br>*Online material* |
| 6 | Topik/ *Topic*:<br>Pangkalan Data Maju<br>*Advance Database*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Makmal (1 jam) / *Lab (1 hour)* | Berman (2013)<br><br>Bahan atas talian<br>*Online material* |
| 7 | Topik/ *Topic*:<br>Teknologi penyimpanan Data Besar: Pengenalan Kepada Pangkalan Data NoSQL (Bahagian 1)<br>*Big data technologies: Introduction to NoSQL databases (Part 1)*<br>Aktiviti / *Activities*:<br>NF2F: Kuliah pra-rakaman (2 jam) / *Pre-recorded Lecture (2 hours)*<br>NF2F: Tutorial pra-rakaman (1 jam) / *Pre-recorded tutorial (1 hour)* | Sabharwal and Edward (2014)<br><br>Bahan atas talian<br>*Online material* |
| 5 May 2025 – 11 May2025 | Cuti pertengahan Semester<br>*Semester Break* | |
| 8 | Topik/ *Topic*:<br>Teknologi penyimpanan Data Besar: Pengenalan Kepada Pangkalan Data NoSQL (Bahagian 2)<br>*Big data technologies: Introduction to NoSQL databases (Part 2)*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (1 jam) / *Lecture (1 hours)*<br>F2F: Tutorial (1 jam) / *tutorial (1 hour)*<br>F2F: Ujian Pertengahan (1 jam, 15%) / *Mid term test (1 hour, 10%)* | Sabharwal and Edward (2014)<br><br>Bahan atas talian<br>*Online material* |
| 9 | Topik/ *Topic*:<br>Teknologi penyimpanan Data Besar: pengaksesan data<br>*Big data technologies: Data access*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Makmal (1 jam) / *Lab (1 hour)* | Bahan atas talian<br>*Online material* |
| 10 | Topik/ *Topic*:<br>Teknologi penyimpanan Data Besar: pengaksesan data<br>*Big data technologies: Data access (Part 2)*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Tutorial (1 jam) / *tutorial (1 hour)* | Bahan atas talian<br>*Online material* |
| 11 | Topik/ *Topic*:<br>Teknologi penyimpanan Data Besar: pengaksesan data<br>*Big data technologies: Data access (Part 3)*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (1 jam) / *Lecture (1 hour)* | Bahan atas talian<br>*Online material* |

# Teaching Plan

| | | |
|---|---|---|
| | F2F: Tutorial (1 jam) / *tutorial (1 hour)*<br>F2F: Ujian Makmal (1 jam, 15%) / *Lab test (1 hour, 15%)* | |
| 12 | Topik/ *Topic*:<br>Asal usul & Kepercayaan terhadap data<br>*Data provenance and data trustworthiness*<br>Aktiviti / *Activities*:<br>F2F: Kuliah (2 jam) / *Lecture (2 hours)*<br>F2F: Makmal (1 jam) / *lab (1 hour)* | Notes<br><br>Bahan atas talian<br>*Online material* |
| 13 | Topik/ *Topic*:<br>Kajian kes<br>*Case study*<br>Aktiviti / *Activities*:<br>F2F: Kuliah atas talian secara langsung (2 jam) / *Live online Lecture (2 hours)*<br>F2F: Makmal atas talian secara langsung (1 jam) / *Live online lab (1 hour)* | Bahan atas talian<br>*Online material* |
| 14 | Topik/ *Topic*:<br>Ulangkaji semua topik<br>*Wrap up of all topics*<br>Aktiviti / *Activities*:<br>NF2F: Forum atas talian, 2.5 jam / *online forum, 2.5 hours*<br>F2F: Laporan dan Pembentangan Projek (0.5 jam, 20%) / *Project Report and Presentation (0.5 hour, 20%)* | Bahan atas talian<br>*Online material*<br><br>Laporan projek<br>*Project report* |

# Assignment & Project

- Topics in Big Data
  - » Agriculture and Crops
  - » Tourism and Hospitality
  - » Veterinary
- Assessment
  - » Assessment 1 - ASSIGNMENT: 15%
  - » Assessment 2 - MID TEST: 10%
  - » Assessment 3 - LAB TEST: 15%
  - » Assessment 4 - PROJECT: 20%
  - » Assessment 5 - ALTERNATIVE 1: 15%
  - » Assessment 6 - ALTERNATIVE 2: 25%

# Things to take note

- Technical Class - Lab Exercise
- Online learning platform:
  - » https://spectrum.um.edu.my/
  - » Microsoft Teams
- Software to be used:
  - » OS: Ubuntu or CentOS, or
  - » Virtualbox (if your system have 8 GB RAM or more)
  - » Hadoop
- Assignment 1!

7 - 13 April 2025 (week 4)

# Online Forum / Discussion

- Join and active participate in the discussions using SPeCTRUM
- Knowledge sharing or anything is allowed (ethically)
- Complete all the assessment

# Q & A

Ts. Dr. Mohd Shahrul Nizam Mohd Danuri
msnizam@um.edu.my | B-3-18
https://www.msnizam.net