Title (Units): **COMP7990 Principles and Practices of Data Analytics (3,2,1)**

Course Aims: This course introduces principal concepts of data management and analysis. It covers various topics including database management, data analytics, data mining, data visualization, and data privacy. It is expected that students can grasp numeracy and practical skills about how to collect, store, analyze, and visualize data.

Prerequisite: Nil

Course Intended Learning Outcomes (CILOs):
Upon successful completion of this course, students should be able to:

No.	Course Intended Learning Outcomes (CILOs)
	Knowledge
1	Describe fundamentals of database management
2	Explain concepts of data analysis techniques and data mining algorithms
3	Describe and explain concepts of data visualization
4	Describe concepts and legal foundations of data security and privacy
	Professional Skill
5	Formulate SQL queries on the database
6	Conduct statistical analysis and design visualization to present analysis results

Calendar Description: This course introduces principal concepts of data management and analysis. Topics mainly include: 1) representation of data in relational database and formulation of data query; 2) knowledge of data analytics and data mining; 3) data visualization concepts; 4) principles of data security and privacy management.

Teaching and Learning Activities (TLAs):

CILOs	Type of TLA
1-4	Students will learn the fundamentals of data management, analysis, and visualization concepts via lectures and tutorials.
5-6	Students will gain practical experiences via laboratory sessions and assignments.

Assessment:

No.	Assessment Methods	Weighting	CILOs to be addressed	Description of Assessment Tasks
1	Continuous Assessment	40%	1-6	Continuous assignments are designed to measure how well students have learnt the fundamentals of data management, analysis, and visualization techniques. Assignments and lab exercises are designed to evaluate students' knowledge and skills. Quizzes are used to determine to what extent the students have achieved the expected learning outcome.
2	Examination	60%	1-4	Final examination questions are designed to identify how far students have achieved intended learning outcomes. Questions will primarily assess students' knowledge in managing, analyzing and visualizing data.

Assessment Rubrics:

Excellent (A)	• Achieve the eight CILOs, demonstrating a mastery of fundamentals of data management, analysis and visualization

	• Have a thorough understanding of concepts and technologies involved, and be able to explain and highlight the key points of these concepts and technologies • Able to perform data storage, analysis and visualization
Good (B)	• Achieve the eight CILOs, demonstrating a good understanding of fundamentals of data management, analysis and visualization • Have a good understanding of concepts and technologies involved, and be able to explain and highlight the key points of these concepts and technologies • Able to perform data storage, analysis and visualization for familiar cases
Satisfactory (C)	• Achieve most of CILOs, demonstrating a basic level of understanding of fundamentals of data management, analysis and visualization • Have a basic understanding of concepts and technologies involved • Able to perform data storage, analysis and visualization for simple cases
Fail (F)	• Achieve less than three of CILOs, having little understanding of fundamentals of data management, analysis and visualization • Have little understanding of concepts and technologies involved, and unable to explain and highlight the key points of these concepts and technologies • Unable to perform data storage, analysis and visualization for simple cases

Course Content and CILOs Mapping:

Content		**CILO No.**
I	Data management	1,5
II	Data analytics	2,6
III	Data mining	2,6
IV	Data visualization	3,6
V	Data security and privacy	4

References:

- Ramez Elmasri and Shamkant B. Navathe. Fundamentals of Database Systems, 7th edition, Addison Wesley, 2015.
- Anil Maheshwari. Data Analytics Made Accessible, Amazon Digital Services LLC, 2023.
- Jure Leskovec, Anand Rajaraman and Jeffrey David Ullman. Mining of Massive Datasets, 3rd edition, Cambridge University Press, 2020.
- Jiawei Han, Micheline Kamber, Jian Pei. Data Mining: Concepts and Techniques, 4th edition, Morgan Kaufmann, 2022.
- Knaflic, Cole N. Storytelling with Data: a Data Visualization Guide for Business Professionals. Hoboken, New Jersey: Wiley, 2015.
- Stephanie Evergreen. Effective Data Visualization: The Right Chart for the Right Data, SAGE Publications, Inc, 2nd edition, 2020.
- Nataraj Venkataramanan and Ashwin Shriram. Data Privacy: Principles and Practice, 1st edition, Chapman and Hall/CRC, 2016.

Course Content:

Topic

I. Data management
   A. Database system concepts
   B. Relational data model
   C. Queries and updates in SQL

II. Data analytics
   A. Statistics concepts
   B. Statistical analysis techniques (such as correlation analysis, linear regression and their applications)
   C. Tools for data analytics

III. Data mining
   A. Data preparation for knowledge discovery
   B. Classification and Clustering algorithms (such as perceptrons, K-NN, K-means, Hierarchical clustering)

IV. Data visualization
   A. Concepts of data visualization
   B. Charts, maps and infographics

V. Data security and privacy
   A. Concepts of data security and privacy
   B. Privacy protection principles