SDSC5003: STORING AND RETRIEVING DATA

Effective Term
Semester A 2025/26

Part I Course Overview

Course Title
Storing and Retrieving Data

Subject Code
SDSC - Data Science
Course Number
5003

Academic Unit
Data Science (DS)

College/School
College of Computing (CC)

Course Duration
One Semester

Credit Units
3

Level
P5, P6 - Postgraduate Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
Nil

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
Nil

Part II Course Details

Abstract

This course offers knowledge of the relational database and an introduction to Hadoop/Spark system including the entity-relationship model for designing the relational database, principles of the database development process, the Structural Query Language for retrieving and storing data via a database, and the introductory level Hadoop/Spark content.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Desribe principles of the relational database design and structural query language syntax	30	x		
2	Explain the efficiency issue in database systems, including storage and indexing as well as query optimization	30	x	x	
3	Discuss modern database techniques such as JSON and NoSQL	10	x	x	
4	Explain the MapReduce computing framework and the basics of Spark	10	x	x	
5	Implement taught knowledge to develop a database application	20		x	x

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lecture	Students will engage in lectures introducing relational database systems, principles of database design and development, syntax of structural query language, as well as fundamentals and introductory applications of Hadoop/Spark	1, 2, 3, 4	26 hours/semester

2	Laboratory work	Students will engage in developing the ability of designing and developing the relational and big database as well as differences between retrieving and storing data via relational and big databases to generate applications.	2, 3, 4, 5	13 hours/semester

Additional Information for LTAs

Lectures cover not only the narrowly focused techniques in engineering economy but also the wider issues of the environment that affect engineering economic decision making. Students are expected to participate in class discussion when needed.

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks ("-" for nil entry)	Allow Use of GenAI?
1	Group Project	2, 3, 4, 5	35	While it's acceptable to seek help from AI tools such as ChatGPT during a group project, avoid outsourcing your work completely. The key is to ensure you fully comprehend the content produced by the AI.	Yes
2	Individual Coursework	1, 2, 3, 4	15	It's okay to use AI tools like ChatGPT for help with your assignments—but don't let AI replace you. Instead, let it assist you. The key is to stay actively involved, critically evaluate the output, and ensure you truly understand the content generated.	Yes

Continuous Assessment (%)

50

Examination (%)

50

Examination Duration (Hours)

2

Minimum Examination Passing Requirement (%)
30

Additional Information for ATs
For a student to pass the course, at least 30% of the maximum mark for the examination should be obtained.

Assessment Rubrics (AR)

Assessment Task
Group Project (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion
35%

Excellent
(A+, A, A-) High

Good
(B+, B, B-) Significant

Fair
(C+, C, C-) Moderate

Marginal
(D) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Individual Coursework (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion
15%

Excellent
(A+, A, A-) High

Good
(B+, B, B-) Significant

Fair
(C+, C, C-) Moderate

Marginal
(D) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task

Examination (for students admitted before Semester A 2022/23 and in Semester A 2024/25 & thereafter)

Criterion
50%

Excellent
(A+, A, A-) High

Good
(B+, B, B-) Significant

Fair
(C+, C, C-) Moderate

Marginal
(D) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Group Project (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion
35%

Excellent
(A+, A, A-) High

Good
(B+, B) Significant

Marginal
(B-, C+, C) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Individual Coursework (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion
15%

Excellent
(A+, A, A-) High

Good
(B+, B) Significant

Marginal
(B-, C+, C) Basic

Failure
(F) Not even reaching marginal levels

Assessment Task
Examination (for students admitted from Semester A 2022/23 to Summer Term 2024)

Criterion
50%

Excellent
(A+, A, A-) High

Good
(B+, B) Significant

Marginal
(B-, C+, C) Basic

Failure
(F) Not even reaching marginal levels

Part III Other Information

Keyword Syllabus
- Introduction of Database and Its Development Process
- Data Modeling (Entity-Relationship model, meta model)
- Database Design Process and development
- Structured Query Language in Relational Database
- Storage and Indexing
- Query Optimization
- JSON and NoSQL
- Fundamentals of MapReduce and Spark
- Applications of Hadoop/Spark

Reading List

Compulsory Readings

	Title
1	Lecture Notes

Additional Readings

	Title
1	NIL