City University of Hong Kong
Course Syllabus

offered by School of Data Science
with effect from Semester A 2022/23

Part I Course Overview

Course Title:	Storing and Retrieving Data
Course Code:	SDSC5003
Course Duration:	One Semester
Credit Units:	3
Level:	P5
Medium of Instruction:	English
Medium of Assessment:	English
Prerequisites: *(Course Code and Title)*	Nil
Precursors: *(Course Code and Title)*	Nil
Equivalent Courses: *(Course Code and Title)*	Nil
Exclusive Courses: *(Course Code and Title)*	Nil

Part II Course Details

1. 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.

2. Course Intended Learning Outcomes (CILOs)

(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs	Weighting (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	Familiarize principles of the relational database design and structural query language syntax	25%	✓		
2.	Understand the efficiency issue in database systems, including storage and indexing as well as query optimization	25%	✓	✓	
3.	Familiarize modern database techniques such as JSON and NoSQL	15%	✓	✓	
4.	Understand the MapReduce computing framework and the basics of Spark	15%	✓	✓	
5.	Implement taught knowledge to develop a database application	20%		✓	✓
		100%			

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 self-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.

3. Teaching and Learning Activities (TLAs)

(TLAs designed to facilitate students' achievement of the CILOs.)

TLA	Brief Description	CILO No.						Hours/week (if applicable)
		1	2	3	4	5		
Lecture	Introduce relational database systems, principles of database design and development, syntax of structural query language, as well as fundamentals and introductory applications of Hadoop/Spark	✓	✓	✓	✓			26 hours/sem
Laboratory work	Assist students to develop 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.		✓	✓	✓	✓		13 hours/sem

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.

4. Assessment Tasks/Activities (ATs)

(ATs are designed to assess how well the students achieve the CILOs.)

Assessment Tasks/Activities	CILO No.						Weighting	Remarks
	1	2	3	4	5			
Continuous Assessment: 50 %								
Group Project		✓	✓	✓	✓		35%	
Individual Coursework	✓	✓	✓	✓			15%	
Examination: 50 % (duration: 2 hours, if applicable)								
Examination	✓	✓	✓		✓		50%	
							100%	

5. Assessment Rubrics

(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Applicable to students admitted in Semester A 2022/23 and thereafter

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B)	Marginal (B-, C+, C)	Failure (F)
1. Group Project	35%	High	Significant	Basic	Not even reaching marginal levels
2. Individual Coursework	15%	High	Significant	Basic	Not even reaching marginal levels
3. Examination	50%	High	Significant	Basic	Not even reaching marginal levels

Applicable to students admitted before Semester A 2022/23

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Group Project	35%	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Individual Coursework	15%	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Examination	50%	High	Significant	Moderate	Basic	Not even reaching marginal levels

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

(An indication of the key topics of the course.)

- 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

2. Reading List

2.1 Compulsory Readings

(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Lecture Notes

2.2 Additional Readings

(Additional references for students to learn to expand their knowledge about the subject.)

NIL