

Lingnan University
Course Syllabus

Course Title	:	Databases and Data Warehouses
Course Code	:	CDS3001
Recommended Study Year	:	3
No. of Credits/Term	:	3
Mode of Tuition	:	Sectional Approach
Class Contact Hours	:	3 hours per week
Category in Major Prog.	:	Required
Prerequisite(s)	:	CDS2003 Data Structures and Object-oriented Programming
Co-requisite	:	Nil
Exclusion	:	Nil
Exemption Requirement	:	Nil

Brief Course Description:

This course is designed to introduce the concepts and principles of database and data warehouse. Various types of databases will be thoroughly discussed in this course, such as objected-oriented, relational, document-oriented, graph, NoSQL, and New SQL. Popular database management systems such as Microsoft SQL Server and/or Oracle will be described and implemented. Topics include data models (ER, relational, and others); query language (Structured Query Language); implementation techniques of database management systems (index structures and query processing); management of semi-structured and complex data; distributed and NoSQL databases; the dimensional modeling technique for designing a data warehouse, and data warehouse architectures, OLAP and the project planning aspects in building a data warehouse.

Aims:

This course aims at introducing the concepts of database and data warehouse. Students will recognize the difference between database and data warehouse. And they will become familiar with various types of database, such as object-oriented, relational, NoSQL and NewSQL. The course will provide an understanding of the database management system, especially how to use the SQL to manipulate the database management system. And how to design and build a data warehouse. A basic understanding and foundation of data science will be provided for students to study further technical skills of data management and business intelligence.

Learning Outcomes (LOs):

On completion of this course, students will be able to:

1. Identify the various types of databases (PLO6);
2. Accurately store different kinds of data to database management systems (PLO6);
3. Implement SQL to query data in database management systems (PLO9);
4. Describe the database and data warehouse design (PLO6);
5. Develop and deploy the database management system to solve some real-world problems (PLO9).
6. Apply essential ethical terminology in data science to clarify the ethical use of data sources, identify potential ethical challenges associated with these sources, and propose improvements to the ethical guidelines that regulate the practices of data scientists.

(PLO1)

Indicative Contents:

Database fundamentals

Overview of databases, such as objected-oriented, relational, document-oriented, graph, NoSQL, and New SQL

The relation model of database normalization

Introduction to Database design

Data modeling and entity-relationship model

Entities, Attributes, and Entity Sets

Relationships and Relationship Sets

Additional Features of the ER Model

Database design and functional dependencies of data attributes sign

Conceptual Design for Large Enterprises

Database Management

Database administration and management

Database processing in applications

SQL: queries, constraints, triggers

The form of a basic SQL query

Union, intersect and except

Nested query, aggregate operators, null values

Complex integrity constraints in SQL

Query development & troubleshooting with Generative AI Tools

Visual Studio Code with Generative AI Extensions such as GitHub Copilot, ChatGPT, Genie AI, etc.

Data warehouse fundamentals

The evaluations of decision support systems

The data warehouse environment

Characteristics of a data warehouse

Data warehouse architectures

Star and snowflake schemas

Data integration and the extraction, transformation, and load (ETL) process

Data warehouse development

Data warehouse design

The data warehouse and data models

The Data Model and Iterative Development

Normalization and Denormalization

Metadata

Complexity of Transformation and Integration

Triggering the Data Warehouse Record

Teaching Method:

There are different teaching and learning activities including lecture and laboratory sections. The concepts and principles of database and data warehouse will be discussed in the lecture sections. The implementation, and manipulation of the database management system by using the SQL will be taught during the laboratory sections. Students are required to perform a group project to apply the concepts and principles covered in this course to critically analyze the given problem(s) and creatively formulate the solution(s). Students will implement the solution(s) using the SQL.

Measurement of Learning Outcomes:

	Class Attendance and Participation	Assignments	Group Project	Examination
1. Identify the various types of databases	X			X
2. Accurately store different kinds of data to database management systems		X	X	
3. Implement SQL to query data in database management systems		X	X	X
4. Describe the database and data warehouse design	X	X	X	X
5. Develop and deploy the database management system to solve some real-world problem			X	
6. Apply essential ethical terminology in data science to clarify the ethical use of data sources, identify potential ethical challenges associated with these sources, and propose improvements to the ethical guidelines that regulate the practices of data scientists.			X	

1. There are a number of classroom activities to evaluate if the students can recognize and identify the various types of database. Students are expected to recognize and describe the database and data warehouse design (LO1, LO4, and PLO6).
2. The assignments require students to accurately store different kinds of data to a database management system. Students are expected to recognize and describe the database and data warehouse design. These assignments evaluate if the student can implement SQL to query data in database management systems (LO2, LO3, L45, PLO6, and PLO9).

3. The Group Project requires students to accurately store different kinds of data to a database management system. Students are expected to recognize and describe the database and data warehouse design. The Group Project evaluates if the student can implement SQL to query data in database management systems. Students need to develop and deploy their database management system to solve some real-world problem (LO2, LO3, LO4, LO5, PLO6, and PLO9).
4. The examination evaluates if the students can recognize and identify the various types of database. It evaluates if the students can recognize and describe the database and data warehouse design. The ability in implementing SQL to query data in database management systems is assessed in the examinations (LO1, LO3, LO4, and PLO6).

Assessment:

Class Attendance and Participation	5%
Assignments	35%
Group Project	20%
<u>Examination</u>	<u>40%</u>
Total	100%

Required/Essential Readings:

1. Kroenke, D. M., D. J. Auer, S. L. Vandenberg, and R. C. Yoder. *Database concepts*. NY, NY: Pearson, 2017.
2. Inmon, W. H. *Building the data warehouse, 4th Edition*. Wiley, 2008.

Recommended/Supplementary Readings:

1. Connolly, Thomas and Carolyn Begg. *Database Systems: A Practical Approach to Design, Implementation, and Management, 6th Edition*. Pearson, 2015.
2. Gillenson, Mark L. *Fundamentals of database management systems 2nd Edition*. Wiley, 2012.
3. Witten, I. H. *Data Mining: Practical Machine Learning Tools and Techniques, 3rd Edition*. Morgan Kaufmann, 2011.
4. Kimball, Ralph, Margy Ross, Warren Thornthwaite, Joy Mundy, and Bob Becker. *The Data Warehouse Lifecycle Toolkit, 2nd Edition*. Wiley, 2008.
5. Hoffer, Jeffrey A., V. Ramesh, and Heikki Topi. *Modern Database Management, 11th Edition*. Pearson 2011.
6. Ponniah, Paulraj. *Data Warehouse Fundamentals for IT Professionals, 2nd Edition*. Wiley 2010.
7. Inmon, W.H., D. Straus, and G. Neushloss. *DW 2.0: The Architecture for the Next Generation of Data Warehousing*. Morgan Kaufmann, 2008.
8. Golfarelli, M., and S. Rizzi. *Data Warehouse Design: Modern Principles and Methodologies*. McGraw-Hill, 2009.
9. Hector, Garcia-Molina, Jeff Ullman, and Jennifer Widom. *Database Systems: The Complete Book, 2nd Edition*. Pearson, 2009.
10. Elmasri, Ramez and Shamkant B. Navathe. *Fundamentals of Database Systems, 7th Edition*. Pearson, 2016.

Important Notes:

- (1) Students are expected to spend a total of 9 hours (i.e. 3 hours of class contact and 6 hours of personal study) per week to achieve the course learning outcomes.
- (2) Students shall be aware of the University regulations about dishonest practice in course work, tests and examinations, and the possible consequences as stipulated in the Regulations Governing University Examinations and Coursework. In particular, plagiarism, being a kind of dishonest practice, is “the presentation of another person’s work without proper acknowledgement of the source, including exact phrases, or summarised ideas, or even footnotes/citations, whether protected by copyright or not, as the student’s own work”. Students are required to strictly follow university regulations governing academic integrity and honesty.
- (3) Students are required to submit writing assignment(s) using Turnitin.
- (4) To enhance students’ understanding of plagiarism, a mini-course “Online Tutorial on Plagiarism Awareness” is available on <https://pla.ln.edu.hk/>.
- (5) Students must adhere to the University’s guidelines and practices when using Generative Artificial Intelligence (GAI) tools. The official documents “*Guidelines for Using GAI Tools at Lingnan University*” and “*Best Practices for Ethical and Responsible Use of GAI Tools in Course Assessments*” can be found here: <https://www.ln.edu.hk/tlc/generative-artificial-intelligence/gai-guidelines-best-practices>.

Rubric for Final Examination of CDS3001 – Databases and Data Warehouses

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Identify the various types of databases	Demonstrates a clear understanding of various types of database. The major difference between most types of databases could be identified in detail.	Demonstrates an understanding of some types of database. The major difference between some types of databases could be identified in detail.	Demonstrates an inadequate level of understanding of many types of database. The major difference between most types of databases cannot be identified in detail.
Implement SQL to query data in database management systems	Demonstrates a good handling of implementation of SQL to query most problems in database management systems. Most results and methods are correct.	Demonstrates a handling of implementation of SQL to query some problems in database management systems. However, a few results and methods are incorrect.	Demonstrates faulty and limited handling of implementation of SQL to query a few problems in database management systems. Moreover, most results and methods are incorrect.
Describe the database and data warehouse design	Demonstrates a clear understanding of the database and data warehouse design. Correctly describes most steps in detail.	Demonstrates an understanding of the database and data warehouse design. The student can describe some steps in specific detail.	Demonstrates an inadequate level of understanding of the database and data warehouse design. Student cannot describe most steps in specific detail.

Rubric for Group Project of CDS3001 – Databases and Data Warehouses

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Accurately store different kinds of data to database management systems	Correctly store many formats of data to database management systems. All/most data are stored in correct formats.	Correctly store some formats of data to database management systems. Yet, some formats of data are applied using a wrong method.	Incorrectly store many formats of data to database management systems. Several formats of data are applied using an incorrect method.
Implement SQL to query data in database management systems	Demonstrate a good handling of implementation of SQL to query all/most problems in database management systems. All/most results and methods are correct.	Demonstrate a handling of implementation of SQL to query some problems in database management systems. However, a few results and methods are incorrect.	Demonstrate a faulty and limited handling of implementation of SQL to query a few problems in database management systems. Moreover, most result and methods are incorrect.
Describe the database and data warehouse design	Demonstrate a clear understanding of the database and data warehouse design. Correctly describe all/most steps in detail.	Demonstrate an understanding of the database and data warehouse design. The student can describe some steps in specific detail.	Demonstrate an inadequate level of understanding of the database and data warehouse design. Student cannot describe most steps in specific detail.
Develop and deploy the database management system to solve some real-world problem	Correctly develop and deploy the database management system to solve some real-world problem. All/most results and methods are correct.	Correctly develop and deploy the database management system to solve some real-world problem. However, some results and methods are incorrect.	Incorrectly develop and deploy the database management system to solve some real-world problem. Most results and methods are incorrect.

Rubric for Assignments of CDS3001 – Databases and Data Warehouses

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Accurately store different kinds of data to database management systems	Correctly stores many formats of data to database management systems.	Correctly store some formats of data to database management systems. Yet, some formats of data are applied using a wrong method.	Incorrectly store many formats of data to database management systems. Several formats of data are applied using an incorrect method.
Implement SQL to query data in database management systems	Demonstrates a good handling of implementation of SQL to query most problems in database management systems. All/most results and methods are correct.	Demonstrates a handling of implementation of SQL to query some problems in database management systems. However, a few results and methods are incorrect.	Demonstrates a faulty and limited handling of implementation of SQL to query a few problems in database management systems. Moreover, most result and methods are incorrect.
Describe the database and data warehouse design	Demonstrates a clear understanding of the database and data warehouse design. Correctly describe all/most steps in detail.	Demonstrates an understanding of the database and data warehouse design. The student can describe some steps in specific detail.	Demonstrates an inadequate level of understanding of the database and data warehouse design. Student cannot describe most steps in specific details.

Rubric for Class Attendance and Participation of CDS3001 – Databases and Data Warehouses

Criteria	Very good (4-6)	Satisfactory (2-4)	Unsatisfactory (0-2)
Identify the various types of databases	Demonstrates a clear understanding of various types of database. The major difference between all/most types of databases could be identified in detail.	Demonstrates an understanding of some types of database. The major difference between some types of databases could be identified in detail.	Demonstrates an inadequate level of understanding of many types of database. The major differences between most types of databases cannot be identified in detail.
Describe the database and data warehouse design	Demonstrates a clear understanding of the database and data warehouse design. Correctly describes all/most steps in detail.	Demonstrates an understanding of the database and data warehouse design. The student can describe some steps in specific details.	Demonstrates an inadequate level of understanding of the database and data warehouse design. Student cannot describe most steps in specific detail.