

BIRLA INSTITUTE OF TECHNOLOGY & SCIENCE, PILANI

WORK INTEGRATED LEARNING PROGRAMMES

COURSE HANDOUT

Part A: Content Design

Course Title	Database Design and Applications
Course No(s)	CSI ZG518 / SS ZG518/SE ZG518
Credit Units	5 (1 unit for lecture, 2 for self-study, 2 for lab / assignment / work integrated activities)
Course Author	R Gururaj
Version No	
Date	

Course Objectives

No	Course Objective
CO1	Enrich the skill and competency of students in Design and Development of Database Systems through experiential learning.
CO2	Provide knowledge about the internals of Database Management Systems.
CO3	Prepare students to tackle the challenges in operationalizing Database Systems, like-maintaining consistency and integrity, performance etc .

Text Book(s)

T1	Ramez Elmasri & Shamkant B. Navathe, Database Systems; Models, Languages, Design and Application Programming, Pearson Education, 7 th Edition, 2017.
----	---

Reference Book(s) & other resources

R1	Abraham Silberschatz, Henry F Korth and S Sudarshan, Database System Concepts, McGraw Hill, 6 th Ed., 2013
R2	Date C.J., An Introduction to Database Systems, Addison Wesley, 8th Ed., 2006.

Content Structure

1. Introduction to Database Management Systems (DBMS) Concepts and Architecture

- 1.1. Database Management Systems Introduction
 - 1.1.1. Basics and advantages of storing data using DBMS
 - 1.1.2. Database systems environment
- 1.2. Database Management System Architecture
 - 1.2.1. Abstraction and Three-schema architecture
 - 1.2.2. Basic modules of DBMS
 - 1.2.3. DBMS Users
2. Data Modeling
 - 2.1. Database Design and ER Modeling
 - 2.1.1. Steps in database Design Process
 - 2.1.2. Concepts and notations
 - 2.1.3. Relationships and constraints
 - 2.1.4. Examples
 - 2.2. Relational Data Model concepts
 - 2.2.1. Relational data model fundamentals
 - 2.2.2. Constraints in Relational data model
 - 2.2.3. Representation of schemas
 - 2.3. ER to Relational Mapping
 - 2.3.1. Mapping rules/guidelines for mapping ER constructs
 - 2.3.2. Mapping rules/guidelines for mapping heirarchies
 - 2.3.3. Examples
3. Relational Query Languages
 - 3.1. Relational Algebra
 - 3.1.1. Basic Relational operations
 - 3.1.2. Other operations and Joins
 - 3.2. SQL
 - 3.2.1. Introduction to SQL
 - 3.2.2. SQL features
 - 3.2.3. SQL join operations
 - 3.2.4. SQL Grouping operations
 - 3.2.5. SQL views
4. Schema Refinement
 - 4.1. Functional Dependencies
 - 4.1.1. Functional dependencies
 - 4.1.2. Inference rules
 - 4.2. Normalization and Decomposition
 - 4.2.1. First and Second Normal forms
 - 4.2.2. Third and BCNF normal forms
 - 4.2.3. Decomposition and desirable properties
 - 4.2.4. Lossless join decomposition
 - 4.2.5. Dependency preserving decomposition
5. Data storage, Hashing and Indexing
 - 5.1. Disk storage
 - 5.1.1. Disk features
 - 5.1.2. Storage capacity
 - 5.1.3. File and Record organization
 - 5.1.4. Types of File organizations
 - 5.1.5. Types of record organizations
 - 5.2. Hashing Techniques

- 5.2.1. Static external Hashing
 - 5.2.2. Dynamic hashing schemes
 - 5.3. Indexing Techniques
 - 5.3.1. Introduction to indexing
 - 5.3.2. Primary and secondary indexing
 - 5.3.3. Multilevel indexing
 - 5.3.4. B+ tree indexing
- 6. Transaction processing, concurrency control and recovery
 - 6.1. Transaction model
 - 6.1.1. Transaction significance
 - 6.1.2. States of a transaction
 - 6.1.3. Schedules- serial and concurrent
 - 6.2. Concurrency Control
 - 6.2.1. Need for Concurrency control
 - 6.2.2. Lock based concurrency control and Deadlocks
 - 6.2.3. Time-stamp based concurrency control
 - 6.3. Database Recovery
 - 6.3.1. Log based recovery techniques
 - 6.3.2. Checkpointing
 - 6.3.3. Shadow paging
- 7. Database Security and Database Programming
 - 7.1. Database security
 - 7.1.1. Introduction to Database Security
 - 7.1.2. Access Control Mechanisms
 - 7.1.3. Statistical Database Security
 - 7.1.4. Flow Control
 - 7.1.5. Other Security Challenges
- 8. Additional Topics
 - 8.1. Query optimization and Tuning
 - 8.1.1. Query execution steps
 - 8.1.2. Query trees
 - 8.1.3. Heuristics
 - 8.1.4. Database tuning concepts
 - 8.2. Database Design methodology and UML
 - 8.2.1. Role of information systems
 - 8.2.2. Design process
 - 8.2.3. UML and tools in Database
 - 8.3. XML data model
 - 8.3.1. Basics of XML
 - 8.3.2. DTD and schemas
 - 8.3.3. XML storage
 - 8.3.4. XML query languages- XPath and XQuery
 - 8.4. Database programming
 - 8.4.1. Accessing databases from programming languages
 - 8.4.2. Triggers
 - 8.4.3. Stored procedures
 - 8.5. Distributed Databases and Client Server Architecture
 - 8.5.1. Purpose of Distributed databases
 - 8.5.2. Managing distributed databases

- 8.5.3. Overview of 3-tier client server architecture
- 8.6. Recent trends in databases
 - 8.6.1. NoSQL Databases
 - 8.6.2. Hadoop

NoSQL Database

<https://www.thoughtworks.com/insights/blog/nosql-databases-overview>

Hadoop

<https://www.mssqltips.com/sqlserverauthor/77/dattatrey-sindol/>

Note: Due to time constraints, all topics listed under module 8 may not be covered. Instructor may decide on some topics from module 8 to be covered as a part of the course.

Learning Outcomes:

No	Learning Outcomes
LO1	More effective and efficient Database Systems leading to more elegant Information System.
LO2	Students are equipped with knowledge on database concepts like- Modelling, Querying, Transactions, Indexing, Recovery etc.
LO3	Students are equipped with the knowledge to tune the Database Systems and Applications for improved performance, response-time etc.

Part B: Contact Session Plan

Academic Term	Second Semester 2020-2021
Course Title	Database Design and Applications
Course No	CSI ZG518 / SS ZG518/SE ZG518
Lead Instructor	

Glossary of Terms:

1. Contact Hour (CH) stands for a hour long live session with students conducted either in a physical classroom or enabled through technology. In this model of instruction, instructor led sessions will be for 22 CH.
 - a. Pre CH = Self Learning done prior to a given contact hour
 - b. During CH = Content to be discussed during the contact hour by the course instructor
 - c. Post CH = Self Learning done post the contact hour
2. Contact Hour (CS) stands for a two-hour long live session with students conducted either in a physical classroom or enabled through technology. In this model of instruction, instructor led sessions will be for 11 CS.
 - a. Pre CS = Self Learning done prior to a given contact session
 - b. During CS = Content to be discussed during the contact session by the course instructor
 - c. Post CS = Self Learning done post the contact session

3. RL stands for Recorded Lecture or Recorded Lesson. It is presented to the student through an online portal. A given RL unfolds as a sequences of video segments interleaved with exercises
4. SS stands for Self-Study to be done as a study of relevant sections from textbooks and reference books. It could also include study of external resources.
5. LE stands for Lab Exercises
6. HW stands for Home Work.
7. M stands for module. Module is a standalone quantum of designed content. A typical course is delivered using a string of modules. M2 means module 2.

Teaching Methodology (Flipped Learning Model)

The pedagogy for this course is centered around flipped learning model in which the traditional class-room instruction is replaced with recorded lectures to be watched at home as per the student's convenience and the erstwhile home-working or tutorials become the focus of classroom contact sessions. Students are expected to finish the home works on time.

Contact Session Plan

- o Each Module (M#) covers an independent topic and module may encompass more than one Recorded Lecture (RL).
- o **Contact Sessions (2hrs each week)** are scheduled alternate weeks after the student watches all Recorded Lectures (RLs) of the specified Modules (listed below) during the previous week
- o In the flipped learning model, Contact Sessions are meant for in-classroom discussions on cases, tutorials/exercises or responding to student's questions/clarification--- may encompass more than one Module/RLs/CS topic.
- o Contact Session topics listed in course structure (numbered CSx.y) may cover several RLs; and as per the pace of instructor/students' learning, the instructor may take up more than one CS topic during each of the below sessions.

Detailed Structure

Introductory Video/Document: << *Introducing the faculty, overview of the course, structure and organization of topics, guidance for navigating the content, and expectations from students*>>

- Each of the sub-modules of **Recorded Lectures** (RLx.y) shall delivered via **30 – 60mins videos** followed by:
- **Contact session** (CSx.y) of 2Hr each for illustrating the concepts discussed in the videos with exercises, tutorials and discussion on case-problems (wherever appropriate); contact sessions (CS) may cover more than one recorded-lecture (RL) videos.

Course Contents

Contact Hour 1

Time	Type	Description	Content Reference
Pre CH	RL1.1	Database Systems DBMS Database System Environment Traditional File Systems Advantages of DBMS over File Systems Disadvantages of DBMS	Ch.1 (T1) 1.1 to 1.6
	RL1.2	DBMS_RL_1.2: Describing and Storing data in DBMS Three Schema Architecture Data Independence Queries Transactions Structure of DBMS Users of DBMS	Ch.2 (T1) 2.1 to 2.4
During CH	CH1	We discuss important concepts like advantages of using DBMS over traditional file systems; Three schema architecture; Data independence; DBMS architecture.	Ch.1 (T1) 1.1 to 1.6 and Ch.2 (T1) 2.1 to 2.4
Post CH	SS1	To be announced at the end of the contact hour (CH1).	
Lab Reference			

Notes: T1 stands for Textbook 1, R1 stand for Reference book 1

Contact Hour 2

Time	Type	Description	Content Reference
Pre CH	RL 2.1	DBMS_RL_2.1: Steps in database Design Process ER Concepts and Notations Class Hierarchies	Ch.3 (T1) 3.1 to 3.7 Ch.4 (T1) 4.1 to 4.3
During CH	CH2	We discuss ER concepts, notations with appropriate examples. We learn how to model databases using ER techniques.	Ch.3 (T1) 3.1 to 3.7 Ch.4 (T1) 4.1 to 4.3
Post CH	HW2	To be announced at the end of the	

		contact hour (CH2).	
Lab Reference			

Contact Hour 3

Time	Type	Description	Content Reference
Pre CH	RL 2.2	DBMS_RL_2.2: Relational data model fundamentals Constraints in Relational data model Representation of schemas	Ch.5 (T1)
During CH	CH3	We discuss to have a better understanding of characteristics and Constraints of Relational model with proper examples.	Ch.5 (T1)
Post CH	SS3	Updates and dealing with constraint violations in Relational model	Ch.5 (T1) 5.3 Ch.2 (R1) 2.6
	QZ3	To be announced at the end of the contact hour (CH3).	
Lab Reference			

Contact Hour 4

Time	Type	Description	Content Reference
Pre CH	RL 2.3	DBMS_RL_2.3: Mapping rules/guidelines for mapping ER constructs Mapping rules/guidelines for mapping hierarchies Examples	Ch.9 (T1) 9.1, 9.2
During CH	CH4	We understand Mapping rules/guidelines for mapping various ER constructs to Relational model with appropriate examples	Ch.9 (T1) 9.1, 9.2
Post CH	HW4	To be announced at the end of the contact hour (CH4).	
Lab Reference			

Contact Hour 5

Time	Type	Description	Content Reference
Pre CH	RL 3.1	DBMS_RL_3.1: Relational Query Languages Formal Query Languages Introduction to relational algebra Relational operators Set Operators Join operator Aggregate functions Grouping Relational Calculus concepts	Ch.8 (T1) 8.1, 8.6
During CH	CH5	We practice writing Relational algebra queries for data retrieval with sample relational schemas. This consolidates our understanding of relational algebra operations.	Ch.8 (T1) 8.1, 8.6
Post CH	SS5	More on Tuple relational calculus (TRC)	Ch.8 (T1) 8.6
	HW5	To be announced at the end of the contact hour (CH5).	
	QZ5	To be announced at the end of the contact hour (CH5).	
Lab Reference			

Contact Hour 6

Time	Type	Description	Content Reference
Pre CH	RL 3.2	DBMS_RL_3.2: Introduction to Structured Query Language (SQL) Features of SQL DDL Statements	Ch.6 (T1) 6.1 to 6.3 Ch.7 (T1) 7.4
During CH	CH6	We practice writing SQL queries for data retrieval with sample relational schemas. This consolidates our understanding of SQL operations.	Ch.6 (T1) 6.1 to 6.3 Ch.7 (T1) 7.4
Post CH	HW6	To be announced at the end of the	

		contact hour (CH6).	
Lab Reference			

Contact Hour 7

Time	Type	Description	Content Reference
Pre CH	RL7	DBMS_RL_3.3 Nested queries and correlated nested queries Use of EXISTS and NOT EXISTS Explicit join operations Aggregate functions Group by and Having clauses Insert / Update / Delete operations Views	Ch. 6(T1) 6.4 Ch.7 (T1) 7.1 to 7.4
During CH	CH7	We practice writing SQL queries for data retrieval with sample relational schemas. This consolidates our understanding of SQL operations.	Ch.6(T1) 6.1 to 6.4 Ch.7 (T1) 7.1 to 7.4
Post CH	SS7	Advanced SQL Concepts: Triggers, Functions, Procedures, and Sequences.	Ch.5 (R1) 5.1 to 5.3
	HW7	To be announced at the end of the contact hour (CH7).	
	QZ7	To be announced at the end of the contact hour (CH7).	
Lab Reference			

Contact Hour 8

Time	Type	Description	Content Reference
Pre CH	RL 4.1	DBMS_RL_4.1: Introduction to Schema Refinement Functional Dependencies Inference Rules Normalization Normal Forms (1NF and 2NF)	Ch.14 (T1) 14.1 to 14.3
During CH	CH8	Understand Functional dependencies, 1 NF and 2 NF with examples	Ch.14 (T1) 14.1 to 14.3 Ch.15(T1) 15.1

Post CH	HW8	To be announced at the end of the contact hour (CH8).	
Lab Reference			

Contact Hour 9

Time	Type	Description	Content Reference
Pre CH	RL9	DBMS_RL_4.2: 3NF and BCNF Decomposition requirements Lossless join decomposition Dependency preserving decomposition Examples	Ch.14 (T1) 14.4 and 14.5 Ch.15 (T1) 15.2 and 15.3
During CH	CH9	Look at 3 NF, BCNF and Decomposition with examples	Ch.14 (T1) 14.4 and 14.5 Ch.15 (T1) 15.2 and 15.3
Post CH	HW9	To be announced at the end of the contact hour (CH9).	
	QZ9	To be announced at the end of the contact hour (CH9).	
Lab Reference			

Contact Hour 10

Time	Type	Description	Content Reference
Pre CH	RL10	DBMS_RL_5.1 Disk pack features Records and Files File operations Ordered and Unordered files	Ch.16 (T1) 16.1 to 16.7
During CH	CH10	To understand Secondary disk storage device; Files, records, blocks on disks	Ch.16 (T1) 16.1 to 16.7
Post CH	SS10	More on Storage devices	Ch.10 (R1) 10.1-10.4
Lab Reference			

Mid semester exams

Contact Hour 11

Time	Type	Description	Content Reference
Pre CH	RL 5.2	DBMS_RL_5.2:	Ch.16 (T1) 16.8

		Introduction to Hashing Internal hashing Collision External hashing Static hashing Dynamic hashing	
During CH	CH11	To understand the concept of Hashing with examples.	Ch.16 (T1) 16.8
Post CH	SS11	More on Hashing	Ch.11 (R1) 11.6 to 11.7
	HW11	To be announced at the end of the contact hour (CH11).	
	QZ11	To be announced at the end of the contact hour (CH11).	
Lab Reference			

Contact Hour 12

Time	Type	Description	Content Reference
Pre CH	RL 5.3	DBMS_RL_5.3: Introduction to Indexing Primary and Secondary indexes Dense and Sparse indexing Multilevel indexing Designing Primary and Multilevel indexes	Ch.17 (T1) 17.1 and 17.2
During CH	CH12	Understanding Primary and Multilevel Indexing with Examples	Ch.17 (T1) 17.1 and 17.2
Post CH	SS12	Other Indexing schemes; Indexing using SQL	Ch.11 (R1) 11.8 to 11.10
Post CH	HW12	To be announced at the end of the contact hour (CH12).	
Lab Reference			

Contact Hour 13

Time	Type	Description	Content Reference
Pre CH	RL 5.4	DBMS_RL_5.4 Introduction to Tree indexing B+ Tree Inserting and Deleting keys into B+	Ch.17 (T1) 17.3

		trees Constructing a B+ tree Designing a B+ tree node structure	
During CH	CH13	Understanding B+ Indexing with Examples	Ch.17 (T1) 17.3
Post CH	SS13	B Tree indexing , Indexing on multiple keys and Other types of Indexing	Ch.17 (T1) 17.3 to 17.5
	HW13	To be announced at the end of the contact hour (CH13).	
Lab Reference			

Contact Hour 14

Time	Type	Description	Content Reference
Pre CH	RL 6.1	DBMS_RL_6.1: Introduction to Transaction Model Significance of Transaction Model States of a transaction ACID properties	Ch.20 (T1) 20.1 to 20.3
During CH	CH14	Discuss on significance of Transaction model and ACID properties in detail with examples	Ch.20 (T1) 20.1 to 20.3
Post CH			
Lab Reference			

Contact Hour 15

Time	Type	Description	Content Reference
Pre CH	RL 6.2	DBMS_RL_6.2 Concurrent Transactions Transaction Schedule Serial and Concurrent Schedules Need for Concurrency Control Conflicting Operations Conflict Equivalent Schedule Test for Conflict Serializability View Equivalent Schedule View Serializability	Ch.20 (T1) 20.4 and 20.5
During CH	CH15	We understand serializability of concurrent schedules with examples	Ch.20(T1) 20.4 and 20.5

Post CH	SS15	View equivalence and view serializability	Ch.20 (T1) Ch.14 (R1) 14.4 to 14.7
Lab Reference			

Contact Hour 16

Time	Type	Description	Content Reference
Pre CH	RL 6.3	DBMS_RL_6.3 Introduction to Concurrency Control Implementing Serializability Lock-based protocols Deadlock condition Two-phase locking protocol Time-stamp based protocols	Ch.21 (T1) 21.1 and 21.2
During CH	CH16	We work with appropriate examples to understand lock-based concurrency control, time-stamp based concurrency control and Deadlock detection technique.	Ch.21 (T1) 21.1 and 21.2
Post CH	SS16	Multiversion Concurrency Control Validation	Ch.21 (T1) 21.3 and 21.4
	HW16	To be announced at the end of the contact hour (CH16).	
	QZ16	To be announced at the end of the contact hour (CH16).	
Lab Reference			

Contact Hour 17

Time	Type	Description	Content Reference
Pre CH	RL 6.4	DBMS_RL_6.4 Introduction to Recovery Recovery strategies Log-based recovery Check-pointing Shadow paging	Ch.22 (T1) 22.1 to 22.4
During CH	CH17	Consolidate our understanding of Recovery concepts with examples.	Ch.22 (T1) 22.1 to 22.4
Post CH	SS17	ARIES; Recovery in multidatabase	Ch.22 (T1)

		Systems; Database backup	
	HW17	To be announced at the end of the contact hour (CH17).	
Lab Reference			

Contact Hour 18

Time	Type	Description	Content Reference
Pre CH	RL 7.1	DBMS_RL_7.1: Introduction to database security; Access control; Statistical database security; Flow control ; Other challenges;	Ch.30 (T1)
During CH	CH18	Look at access control options in Oracle and understand statistical databases.	
Post CH			
Lab Reference			

Contact Hour 19, 20 and 21

Some topics from module 8 to be covered as decided by the instructor.

Contact Hour 22

Revision Lecture

Detailed Plan for Lab work

Install SQLite & SQLiteBrowser on your laptop for carrying out the lab exercises

Lab No	Topic	Lab Objective
1	Basics of SQL	Understand installation of DBMS such as SQLite and SQLiteBrowser and explore its features
2	Data Definition Language (DDL)	Understand how to CREATE tables
4	Data Manipulation Language(DML)	Basics of SQL: SELECT & JOIN
5	DML	Understand usage of Aggregate functions
6	DML	Understand how to use UPDATE and DELETE operations

7.	PLSQL	Writing Triggers, Functions, Stored Procedures
----	-------	--

Please refer to the attached lab sheet to perform the SQL exercises

Assignment

1. Think of a web-based software application (with a central database), that you think will be useful to many people (1 week)
2. Write a brief description of the application (1 week)
 - Who will be the users?
 - What are the benefits of this application?
 - List of functions & features of the application
 - How many users will use it simultaneously?
3. Draw ER diagram for the application(2 weeks)
4. Convert ER model to relational model (2 weeks)
5. Write SQL queries to fulfil the end user needs (2 weeks)
6. Design indices to enhance the performance of queries and justify your choice of indices

Work Integrated activities (8 weeks)

Activity No	Topic	Description
1	Conceptual model	Draw ER diagram of the database used by your application
2	Logical database model	Study the tables of your application & determine if it satisfies 3 rd Normal Form. Recommend improvements if any, to the table design
3	Physical design	Study the indexes of your application & suggest improvements if any.
4	Database administration	Talk to your DBA and find out what their challenges are. Think of ways to address these challenges and discuss your suggestions with them. What is the outcome of these discussions?
5	Security	What security features of DBMS are implemented in your application (set of applications). Give examples of situations where these features are used.

Evaluation Scheme:

Legend: EC = Evaluation Component; AN = After Noon Session; FN = Fore Noon Session

No	Name	Type	Duration	Weight	Day, Date, Session, Time
EC-1	Quiz-I	Online	-	5%	February 1-15, 2021
	Quiz-II	Online	-	5%	March 1-15, 2021
	Quiz – III	Online	-	5%	April 1-15, 2021
	Lab / Assignment	Online		10%	To be announced

EC-2	Mid-Semester Test	Closed Book	2 hours	30%	Sunday, 07/03/2021 (AN) 2 PM – 4 PM
EC-3	Comprehensive Exam	Open Book	3 hours	45%	Sunday, 02/05/2021 (AN) 2 PM – 5 PM

Note - Evaluation components can be tailored depending on the proposed model.

Important Information:

Syllabus for Mid-Semester Test (Closed Book): Topics in CS 1-5.

Syllabus for Comprehensive Exam (Open Book): All topics given in plan of study

Evaluation Guidelines:

1. For Closed Book tests: No books or reference material of any kind will be permitted. Laptops/Mobiles of any kind are not allowed. Exchange of any material is not allowed.
2. For Open Book exams: Use of prescribed and reference text books, in original (not photocopies) is permitted. Class notes/slides as reference material in filed or bound form is permitted. However, loose sheets of paper will not be allowed. Use of calculators is permitted in all exams. Laptops/Mobiles of any kind are not allowed. Exchange of any material is not allowed.
3. If a student is unable to appear for the Regular Test/Exam due to genuine exigencies, the student should follow the procedure to apply for the Make-Up Test/Exam. The genuineness of the reason for absence in the Regular Exam shall be assessed prior to giving permission to appear for the Make-up Exam. Make-Up Test/Exam will be conducted only at selected exam centres on the dates to be announced later.

It shall be the responsibility of the individual student to be regular in maintaining the self-study schedule as given in the course handout, attend the lectures, and take all the prescribed evaluation components such as Assignment/Quiz, Mid-Semester Test and Comprehensive Exam according to the evaluation scheme provided in the handout.

Appendix

Lab Exercise in SQL

1. Install SQLite & SQLiteBrowser on your laptop (free database)
2. Explore the features of SQLiteBrowser
3. Create a database with the schema given at the end of this document
4. Populate these tables of this schema with sample data
5. Write & execute SQL statements to obtain the following information from the database

Section 1

1. Get details of employees whose salary is greater than or equal to Rs. 30 lakhs
2. Get details of employees working in 'Research' department.
3. For all employees, get their SSN, their name and name of the project they are working on
4. For all employees, get their SSN, their name and the name of their manager (supervisor)

Section 2

1. Find total number of employees
2. Find total number of employees in 'Research' department
3. Find the number of projects controlled by each department (get department #,

department name & # of projects controlled)

Section 3

1. Find departments which control more than 5 projects
2. Get Name & SSN of Employees with more than 2 dependents

Section 4

1. Increase salary of employees working in 'HR' department by 10%
2. Delete project whose project number = 12345

Database schema

