c. Compare density distribution for features age and passenger fare

d. Use a pair plot to show pairwise bivariate distribution

4. Using Titanic dataset, do the following

a. Find total number of passengers with age less than 30

b. Find total fare paid by passengers of first class

c. Compare number of survivors of each passenger class

5. Download any dataset and do the following

a. Count number of categorical and numeric features

b. Remove one correlated attribute (if any)

c. Display five-number summary of each attribute and show it visually

Project: Students are encouraged to work on a good dataset in consultation with their faculty and apply the concepts learned in the course.

Note: Examination scheme and mode shall be as prescribed by the Examination Branch, University of Delhi, from time to time.

DSC11/DSC05/GE3a: DATABASE MANAGEMENT SYSTEMS

Credit distribution, Eligibility and Prerequisites of the Course

Course title & Code	Credits	Credit distribution of the course			Eligibility criteria	Pre-requisite of the course (if any)
		Lecture	Tutorial	Practical/ Practice		
Database Management Systems	4	3	0	1	Pass in Class XII	NIL

Course Objectives

The course introduces the students to the fundamentals of database management system and its architecture. Emphasis is given on the popular relational database system including data

models and data manipulation. Students will learn about the importance of database structure and its designing using conceptual approach using Entity Relationship Model and formal approach using Normalization. The importance of file indexing and controlled execution of transactions will be taught. The course would give students hands-on practice of structured query language in a relational database management system and glimpse of basic database administration commands.

Learning outcomes

On successful completion of the course, students will be able to:

● Use database management system software to create and manipulate the database.

● Create conceptual data models using entity relationship diagrams for modeling real-life situations and designing the database schema.

● Use the concept of functional dependencies to remove redundancy and update anomalies.

● Apply normalization theory to get a normalized database scheme.

● Write queries using relational algebra, a procedural language.

Syllabus

Unit 1   **(5 hours)**

Introduction to Database: Purpose of database system, Characteristics of database approach, data models, database management system, database system architecture, three-schema architecture, components of DBMS, data independence, and file system approach vs database system approach.

Unit 2   **(7 hours)**

Entity Relationship Modeling: Conceptual data modeling - motivation, entities, entity types, attributes, relationships, relationship types, constraints on relationship, Entity Relationship diagram notation.

Unit 3   **(7 hours)**

Relational Data Model: Update anomalies, Relational Data Model - Concept of relations, schema-instance distinction, keys, relational integrity constraints, referential integrity and foreign keys, relational algebra operators and queries.

Unit 4   **(12 hours)**

Structured Query Language (SQL): Querying in SQL, DDL to create database and tables, table constraints, update database-update behaviors, DML, aggregation functions group by and having clauses, retrieve data from the database, generate and query views. Access and manipulate databases using ODBC. Basic Database administration SQL commands.

Unit 5 **(10 hours)**

Database Design: Mapping an Entity Relationship model to relational database, functional dependencies and Normal forms, 1NF, 2NF, 3NF and BCNF decompositions and desirable properties of them.

Unit 6 **(4 hours)**

Data Storage and Indexes: Need of file indexes, file organizations, index structures, single- and multi-level indexing, concurrent execution of transactions, ACID properties,.

Essential/recommended readings

1. Elmasri, R., Navathe, B. S. Fundamentals of Database Systems, 7th Edition, Pearson Education, 2015.

2. Krogh, J. W. MySQL Connector/Python Revealed: SQL and NoSQL Data Storage Using MySQL for Python Programmers, Apress, 2018.

3. Murach J. Murach's MySQL, 3rd edition, Pearson, 2019.

Additional References

1. Ramakrishnan, R., Gehrke J. Database Management Systems, 3rd Edition, McGraw Hill, 2014.

2. Silberschatz, A., Korth, H. F., Sudarshan S. Database System Concepts, 7th Edition, McGraw Hill, 2019.

3. Connolly, T. M., Begg, C. E. Database Systems: A Practical Approach to Design, Implementation, and Management, 6th edition, Pearson, 2019.

Practicals (30 hours)

Create and use the following student-course database schema for a college to answer the given queries using the standalone SQL editor.

Here, Rollno (ADMISSION) and SID (ADMISSION) are foreign keys. Note that course type may have two values viz. Fulltime and Parttime and a student may enroll in any number of courses

1. Retrieve names of students enrolled in any course.
2. Retrieve names of students enrolled in at least one part time course.
3. Retrieve students' names starting with letter 'A'.
4. Retrieve students' details studying in courses 'computer science' or 'chemistry'.
5. Retrieve students' names whose roll no either starts with 'X' or 'Z' and ends with '9'
6. Find course details with more than N students enrolled where N is to be input by the user.
7. Update student table for modifying a student name.
8. Find course names in which more than five students have enrolled
9. Find the name of youngest student enrolled in course 'BSc(P)CS'
10. Find the name of most popular society (on the basis of enrolled students)
11. Find the name of two popular part time courses (on the basis of enrolled students)
12. Find the student names who are admitted to full time courses only.
13. Find course names in which more than 30 students took admission
14. Find names of all students who took admission to any course and course names in which at least one student has enrolled
15. Find course names such that its teacher-in-charge has a name with 'Gupta' in it and the course is full time.
16. Find the course names in which the number of enrolled students is only 10% of its total seats.
17. Display the vacant seats for each course
18. Increment Total Seats of each course by 10%
19. Add enrollment fees paid ('yes'/'No') field in the enrollment table.
20. Update the date of admission for all the courses by 1 year.
21. Create a view to keep track of course names with the total number of students enrolled in it.
22. Count the number of courses with more than 5 students enrolled for each type of course.
23. Add column Mobile number in student table with default value '9999999999'
24. Find the total number of students whose age is > 18 years.
25. Find names of students who are born in 2001 and are admitted to at least one part time course.

Create and use the following student-society database schema for a college to answer the given (sample) queries using the standalone SQL editor.

II. Do the following database administration commands:

Create user, create role, grant privileges to a role, revoke privileges from a role, create index