

Course Code	Course Title				Category
	Database Management Systems				
Contact Hours per Week			CA	FE	Credits
L	T	D/P			
4	0	4	40	60	3
Prerequisite: <ul style="list-style-type: none"> Fundamentals of data structures Basic Knowledge of file handling 					
Course Objectives: <ul style="list-style-type: none"> To learn the concepts of database management and querying the databases. To be familiar with various database design techniques and practice. To understand various transaction processing and concurrency control. To study the concepts of distributed databases and their management. To understand technical concepts to handle big data with distributed databases. 					

COURSE CONTENT

Unit I – INTRODUCTION OF DATABASES and SQL

09

Database Concepts, Three-schema architecture of a database, Data Models ER model, Relational Model, ER to Table Conversion. Relational Algebra: Select, Project, Union, Set difference, Join, SQL-Characteristics and advantages, SQL Data Types and Literals, DDL, DML, DCL, TCL, Views, Indexes. PLSQL : Concept of Stored Procedures, Functions, Cursors, Triggers. NOSQL- MongoDB CRUD Operations, SQL VsNoSQL Databases, Introduction of databricks,

Unit II – DATABASE DESIGN

09

Functional Dependency, Purpose of Normalization, Data Redundancy and Update Anomalies, Functional Dependency Single Valued Dependencies. Single Valued Normalization: 1NF, 2NF, 3NF, BCNF. Decomposition: lossless join decomposition and dependency preservation, Multi valued Normalization (4NF), Join Dependencies and the Fifth Normal Form

Unit III -DATABASE STORAGE, PROCESSING AND TRANSACTION

09

Query processing and query optimization, Basic concept of a Transaction, Transaction Management, ACID Properties of Transactions, Concept of Schedule, Serial and Concurrent Schedule, Serializability: Conflict and View, Cascaded Aborts, Recoverable and Non-recoverable Schedules, Concurrency Control: Need, Locking based Protocol, Deadlocks-Prevention, Detection Techniques, Recovery methods : Shadow Paging and Log Based Recovery, Checkpoints, Introduction to RDD,RDD operations.

Unit IV – ADVANCE CONCEPTS OF DATABASES**09**

Database Architectures: Centralized and ClientServer Architectures, Database Connectivity using Java/Python with SQL and NoSQL databases. Introduction to Parallel Databases, Architecture of Parallel Databases. Introduction to Distributed Databases, Distributed Transactions. 2PC, 3PC protocols, Introduction to Data Mining and clustering.

Unit V - EMERGING DATA HANDLING TECHNIQUES**09**

Introduction to Big data, Handling large datasets using Map-Reduce and Hadoop, Parquet file Format, Introduction to Hbase data model and hbase region. Introduction to emerging database technologies- Cloud Databases, Mobile Databases, SQLite Database, XML Databases, Introduction of Apache spark, Features and uses of Apache spark

Course Outcomes: (Students will be able to-)

- Apply the concepts of database design and SQL.
 - Query a database using SQL, PL/SQL and NoSQL commands.
 - Design and implement a big data store using HBase.
 - Analyze big data using map-reduce programming.
-

Text Books:

1. Silberschatz A., Korth H., Sudarshan S., "Database System Concepts", 6 th Edition, McGraw Hill Publishers, ISBN 007120413X
2. Connally T., Begg C., "Database Systems", 3rd Edition, Pearson Education, 2002, ISBN 8178088614
3. "MongoDB: The Definitive Guide" by Kristina Chodorow, O'Reilly Publications
4. "Principles of Distributed Database Systems", by M. Tamer Özsu, Patrick Valduriez, Springer

References:

1. Data Mining: Concepts and Techniques, Jiawei Han, Micheline Kamber, Jian Pei, Elsevier
2. Big Data: Understanding How Data Powers Big Business, Bill Schmarzo, Wiley
3. Hadoop: The Definitive Guide, Fourth Edition, Tom White, O'Reilly
4. HBase: The Definitive Guide, Fourth Edition, Lars George, O'Reilly