SEMESTER – 2

20MCA102	ADVANCED DATABASE MANAGEMENT YSTEMS	CATEGORY	L	T	P	CREDIT
		GENERAL	3	1	0	4

Preamble: This course provides the basic concepts and terminology related to relational and non-relational database management systems. The concept of advanced DBMS techniques and new generation databases like MongoDB, HBase and Cassandra are also introduced. This course serves as a prerequisite for many advanced courses in Data Science and Machine Learning areas.

Prerequisite: Basic knowledge in Database Management Systems.

Course Outcomes: After the completion of the course the student will be able to

CO 1	Understand the fundamentals of relational database systems including: data models, database architectures and ER features.
CO 2	Analyze and apply the different normalization techniques.
CO 3	Assess the basic issues of transaction processing and concurrency control.
CO 4	Understand the roles that databases play in organizations and familiarize with basic database storage, file organization, database accessing techniques.
CO 5	Understand the basics of query processing, object-oriented, distributed databases.
CO 6	Analyze non-relational database systems and structures and XML.

Mapping of course outcomes with program outcomes

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12
CO 1				1			1					
CO 2	3	3	3	2			2	2			2	2
CO 3	1	2	2	2		2					2	2
CO 4					1		1					
CO 5	1			1								
CO 6	1											

Assessment Pattern

Bloom's Category	Continuous Assessment Tests		End Semester Examination
	1	2	
Remember(K1)	10	10	10
Understand(K2)	20	20	20
Apply(K3)	20	20	30
Analyse(K4)			
Evaluate(K5)			
Create(K6)			

Mark distribution

Total Marks	CIE	ESE	ESE Duration
100	40	60	3 hours

Continuous Internal Evaluation Pattern:

Attendance : 8 marks

Continuous Assessment Test (2 numbers) : 20 marks

Assignment/Quiz/Course project : 12 marks

End Semester Examination Pattern: There will be two parts; Part A and Part B. Part A contains 10 compulsory short answer questions, 2 from each module. Each question carries 3 marks. Part B contains 2 questions from each module of which student should answer any one. Each question can have maximum 2 sub-divisions and carry 6 marks.

Course Level Assessment Questions

Course Outcome 1 (CO1):

1. Examine why databases are important. (K3)
2. Describe the basic features of the relational data model and discuss their importance to the end user and the designer. (K2)
3. Analyze the graphic depiction of relationships among the entities and examine how these depictions help in the database design process. (K3 & K4)

Course Outcome 2 (CO2):

1. Evaluate and design good table structures to control data redundancies and anomalies. (K5 & K6)

Course Outcome 3(CO3):

1. Explain the database transaction and its properties. (K2)
2. Describe concurrency control and analyze the role it plays in maintaining the database integrity. (K2 & K4)
3. Assess the common algorithms for concurrency control. (K5)
4. Define deadlock and discuss the strategies for managing deadlocks. (K1 & K2)
5. Examine how database recovery management is used to maintain database integrity. (K3)

Course Outcome 4 (CO4):

1. Discuss the various disk-organization techniques. (K2)
2. Describe the various data structures that allow fast access to data. (K2)
3. Analyze and examine the different indexing techniques. (K3 & K4)

Course Outcome 5 (CO5):

1. Describe the basics of query processing and evaluate the query processing cost. (K2 & K5)
2. Analyze the concept of object oriented databases and distributed databases. (K4)

Course Outcome 6 (CO6):

1. Explain the concept of XML. (K2)
2. Describe the various NoSQL databases. (K2)

Model Question Paper

Reg No.:_____		Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

MODEL QUESTION PAPER M.C.A.DEGREE EXAMINATION

Course Code: 20MCA102

Course Name: ADVANCED DATABASE MANAGEMENT SYSTEMS

Max. Marks: 60		Duration: 3 Hours

PART A

		Answer all questions, each carries3 marks.	Marks
1		Define weak entity set with an example.	(3)
2		With the help of a diagram explain the different levels of data abstraction?	(3)
3		Differentiate between BCNF and 3NF with an example.	(3)
4		Explain functional dependency with suitable example.	(3)
5		Discuss the ACID properties of transaction.	(3)
6		Define deadlock and discuss the strategies for managing deadlocks.	(3)
7		Diagrammatically represent the basic steps in query processing.	(3)
8		Differentiate static and dynamic hashing.	(3)
9		Illustrate the different types of Distributed Databases.	(3)
10		Define collection and document in MongoDB.	(3)

PART B

Answer any one question from each module. Each question carries 6 marks.

Module I

11		Draw an E-R diagram of a college database with entities student, staff, course, teacher, clerk, department & hostel? Relationship names must be meaningful	(6)

		and there should be an ISA relationship also in diagram.	
		OR	
12		Explain the relational model concept and discuss the different relational model constraints.	(6)
		Module II	
13		Analyse the common anomalies found in databases? How can we eliminate it through normalization?	(6)
		OR	
14		Define Normalization. Explain 1NF, 2NF and 3NF in detail.	(6)
		Module III	
15		Define deadlock. How can we deal with deadlocks?	(6)
		OR	
16		Explain concurrency control with locking methods.	(6)
		Module IV	
17		Explain the various RAID levels with appropriate diagrams.	(6)
		OR	
18		Differentiate between Dense index and Sparse index with example.	(6)
		Module V	
19		Explain HBase and Cassandra.	(6)
		OR	
20		Explain XML and its applications.	(6)
		****	

SYLLABUS

Module I:

Relational Databases:- Introduction - Purpose of Database System – Database System Applications - View of data: Data Abstraction, Instances and Schemas, Data Models – Database Architecture - Database Users and Administrators: Database Users and Interfaces, DBA – Introduction to the Relational Model: Structure of Relational Database, database Schema, Keys, Relational Query language – The Relational Algebra: Fundamental Operations, Formal definition of the relational algebra, additional relational algebra operations – The Entity-Relationship model: Entity Set, Relationship Set, Attributes – Constraints: Mapping cardinalities, Key Constraints, Participation Constraints - E-R Diagrams: Basic structure, Complex attributes, Roles, Non binary relationship sets, Weak Entity Set, Relational Database Design using ER- to Relational Mapping – Extended ER Features: Specialization, Generalization, Attribute inheritance, Constraints on generalization, Aggregation.

Module II:

Database Design:- Database Tables and Normalization – The Need for Normalization – The Normalization Process: Inference Rules for Functional Dependencies (proof not needed) - Minimal set of Functional Dependencies - Conversion to First Normal Form, Conversion to Second Normal Form, Conversion to Third Normal Form - Improving the Design - Surrogate Key Considerations - Higher Level Normal Forms: Boyce/Codd Normal Form, Fourth Normal Form, Join dependencies and Fifth Normal Form – Normalization and Database Design.

Module III:

Transaction Management and Concurrency Control:- Transaction: Evaluating Transaction Results, Transaction Properties, Transaction Management with SQL, The Transaction Log – Concurrency Control: Lost Updates, Uncommitted Data, Inconsistent Retrievals, The Scheduler – Concurrency Control with Locking Methods: Lock Granularity, Lock Types, Two Phase Locking to Ensure Serializability, Deadlocks – Concurrency Control with Timestamping Methods: Wait/Die and Wait/Wound Schemes – Concurrency Control with Optimistic Methods - Database Recovery Management: Transaction Recovery.

Module IV:

Data Storage and Querying:- RAID – File Organization – Organization of Records in Files – Indexing and Hashing: Basic concept, Ordered Indices, B+ tree Index Files: Structure of a B+-Tree (structure only, algorithms not needed) - B tree index files – Static Hashing – Dynamic Hashing – Query Processing: Overview - Selection Operation.

Module V:

System Architecture, Object Oriented Databases, XML and NoSQL:- Distributed Databases: Homogeneous and Heterogeneous Databases, Distributed Data Storage,

Distributed Transactions - Object Based Databases: Overview, Complex Data types, Structured types and inheritance in SQL, Table Inheritance, Array and Multiset types in SQL, Object identity and reference types in SQL - XML: DTD and XML Schema, XML presentation, XML Applications - Next Generation Databases: Distributed Relational Databases - Nonrelational Distributed Databases - MongoDB Sharding and Replication - Hbase - Cassandra - CAP Theorem.

Text Books

1. Abraham Silberschatz, Henry F. Korth, S. Sudarshan," ***Database System Concepts***", McGraw Hill Education, 6th Edition, 2011. *(for Module 1 Refer Chapter 1 [1.1 to 1.3, 1.9,1.12], Chapter 2 [2.1-2.3,2.5], Chapter 6 [6.1], Chapter 7 [7.2, 7.3, 7.8(7.81. To 7.8.5)], for Module 4 Refer Chapter 10 [10.3, 10.5, 10.6], Chapter 11 [11.1, 11.2, 11.3(11.3.1), 11.4.5 and module 5 Refer Chapter 19 [19.1,19.2, 19.3 - Distributed Databases], Refer Chapter 22 [22.1 to 22.6 - Object Based Databases]).*
2. Ramez Elmasri, Shamkant B.Navathe, " ***Fundamentals of Database Systems*** ", Pearson Education, 5th Edition, 2007. *(for Module 1 - Refer Chapter 7 [7.1] - 7.1.1 - Relational Database Design using ER- to Relational Mapping]) and for Module 2 - Refer Chapter 10 [10.2.2 and 10.2.4], Refer Chapter 11 [11.4 - Join dependencies and Fifth Normal Form).*
3. Guy Harrison, "***Next Generation Databases: NoSQL, NewSQL, and Big Data***", Apress, 1st Edition, 14 December 2015. *Refer Chapters 8 and 3 (for Module 5 - Next Generation Databases and CAP Theorem).*

4. Rob, Peter and Carlos Coronel, "***Database Principles: Fundamentals of Design, Implementation and Management***", 9th Edition, 2011. *(for Module 2, refer chapter 6) and (for module 3, refer chapter 10) and (for Module 5, refer Chapter 14 -XML).*

Reference Books

1. Ashutosh Kumar Dubay, "***Database Management Concepts***", S.K. Kataria & Sons, 1st Edition (2012).
2. Raghu Ramakrishnan and Johannes Gehrke, "***Database Management Systems***", McGraw Hill, 3rd Edition (2014).
3. Thomas M Connolly and Carolyn E Begg, "***Database systems- A Practical Approach to Design, Implementation and Management***", Pearson Education, 4th Edition (2014).

Web Resources

1. Introduction to Databases (nptel) https://nptel.ac.in/courses/106/106/106106220/
2. Database Design (nptel) https://nptel.ac.in/courses/106/106/106106093/
3. Introduction to Database Systems and Design https://nptel.ac.in/courses/106/106/106106095/
4. Fundamentals of Database Systems https://nptel.ac.in/courses/106/104/106104135/#
5. Database Management Essentials (Coursera) https://www.coursera.org/learn/database-management
6. Database Systems Concepts & Design https://www.udacity.com/course/database-systems-concepts-design--ud150

Course Contents and Lecture Schedule

No.	Topic	No. of Lectures
1	***Module I: Relational Databases***	***15 hrs***
1.1	Introduction - Purpose of Database System - Database System Applications	1 hr
1.2	View of data: Data Abstraction, Instances and Schemas, Data Models	1 hr
1.3	Database Architecture	1 hr
1.4	Database Users and Administrators: Database Users and Interfaces, DBA	1 hr
1.5	Introduction to the Relational Model: Structure of Relational Database, database Schema, Keys, Relational Query language	1 hr
1.6	The Relational Algebra: Fundamental Operations, Formal definition of the relational algebra, additional relational algebra operations	2 hr
1.7	The Entity-Relationship model: Entity Set, Relationship Set, Attributes	1 hr
1.8	Constraints: Mapping cardinalities, Key Constraints, Participation Constraints	2 hr
1.9	E-R Diagrams: Basic structure, Complex attributes, Roles, Non binary relationship sets, Weak Entity Set	1 hr
1.10	Relational Database Design using ER- to Relational Mapping	2 hr
1.11	Extended ER Features: Specialization,Generalization, Attribute inheritance, Constraints on generalization, Aggregation.	2 hr
2	***Module II: Database Design***	***9 hrs***
2.1	Database Tables and Normalization - The Need for Normalization	1 hr
2.2	The Normalization Process: Inference Rules for Functional Dependencies (proof not needed) - Minimal set of Functional Dependencies - Conversion to First Normal Form, Conversion to Second Normal Form	2 hr
2.3	Conversion to Third Normal Form	1 hr
2.4	Improving the Design - Surrogate Key Considerations	1 hr

No	Topic	No. of Lectures
2.5	Higher Level Normal Forms: Boyce/Codd Normal Form	1 hr
2.6	Fourth Normal Form	1 hr
2.7	Join dependencies and Fifth Normal Form	1 hr
2.8	Normalization and Database Design	1 hr
3	***Module III: Transaction Management and Concurrency Control***	***9 hrs***
3.1	Transaction: Evaluating Transaction Results, Transaction Properties	1 hr
3.2	Transaction Management with SQL, The Transaction Log	1 hr
3.3	Concurrency Control: Lost Updates, Uncommitted Data, Inconsistent Retrievals, The Scheduler	2 hr
3.4	Concurrency Control with Locking Methods: Lock Granularity	1 hr
3.5	Lock Types, Two Phase Locking to Ensure Serializability	1 hr
3.6	Deadlocks	1 hr
3.7	Concurrency Control with Timestamping Methods: Wait/Die and Wait/Wound Schemes, Concurrency Control with Optimistic Methods, Database Recovery Management: Transaction Recovery	2 hr
4	***Module IV: Data Storage and Querying***	***10 hrs***
4.1	RAID	1 hr
4.2	File Organization	1 hr
4.3	Organization of Records in Files	1 hr
4.4	Indexing and Hashing: Basic concept, Ordered Indices	1 hr
4.5	B+ tree Index Files: Structure of a B+-Tree, B tree Index Files	2 hr
4.6	Static Hashing, Dynamic Hashing	2 hr
4.7	Query Processing: Overview, Selection Operation	2 hr

No	Topic	No. of Lectures
5	***Module V: System Architecture, Object Oriented Databases, XML and NoSQL***	***13 hrs***
5.1	Distributed Databases: Homogeneous and Heterogeneous Databases, Distributed Data Storage, Distributed Transactions	2 hr
5.2	Object Based Databases: Overview, Complex Data types	1 hr
5.3	Structured types and inheritance in SQL	1 hr
5.4	Table Inheritance	1 hr
5.5	Array and Multiset types in SQL	1 hr
5.6	Object identity and reference types in SQL	1 hr
5.7	XML: DTD and XML Schema	1 hr
5.8	XML presentation, XML Applications	1 hr
5.9	Next Generation Databases: Distributed Relational Databases - CAP Theorem	1 hr
5.10	Norelational Databases – MongoDB Sharding and Replication	1 hr
5.11	Hbase	1 hr
5.12	Cassandra	1 hr