KONERU LAKSHMAIAH EDUCATION FOUNDATION
(Deemed to be University estd, u/s, 3 of the UGC Act, 1956)
(NAAC Accredited "A++" Grade University)
Green Fields, Guntur District, A.P., India – 522502
Department of Electronics and Communication Engineering
(DST - FIST Sponsored Department)

Active Learning Method

Program: B. Tech Academic Year / Yr-Sem : 2024 - 25 / II - II Sem

Course Title & Code: **DBMS & 23AD2102R**

Date: **10.04.2025** Time: Venue:

CO#	4
Topics	DSPF(Hadoop)
Type of ALM	Case Study
Learning Approach	Participatory Learning

Activity:

Task:

1. **Explain how you would design a data storage solution** for a company that collects daily transaction data from multiple locations. Outline how data would be partitioned and managed across nodes.

Data Storage Design for Multi-location Daily Transactions

1. **Architecture:** Use a distributed database (e.g., Cassandra, DynamoDB) or data lake (e.g., S3 + Athena) for scalability and high availability.
2. **Partitioning:**
   - **Primary key:** `Location ID`
   - **Secondary key:** `Date`
   - Distribute data across nodes by location or time (e.g., monthly partitions).
3. **Data Ingestion:**
   - Stream data via Kafka/Kinesis.
   - Write to nodes based on partitioning logic.
4. **Replication:** Store 2–3 replicas per partition across different nodes for fault tolerance.
5. **Cold Data Management:**
   - Archive older data to cheaper storage (e.g., S3 Glacier).
   - Apply lifecycle rules.
6. **Query Optimization:**
   - Index by date, location.
   - Use materialized views for fast reporting.
7. **Monitoring & Scaling:**
   - Monitor performance.
   - Auto-scale storage and compute resources.