

The BDT exam will contain 4 questions. Q1 is 34 marks the rest of the q's are 33 marks; you do 3 questions. Make sure to attempt all 3. You can complete all 4 questions, I will take your best 3.

Object relational	Object Relational DBMS <ul style="list-style-type: none"> • Object Types (ADT) • Inheritance • Columns Objects and Row Objects • Constraint Definition • Use of REF, SCOPE REF, Dangling REF; • Collection Types • Nested Table Type; VARRAY
NoSQL	<ul style="list-style-type: none"> • Domain Driven Design using Aggregates in the context of NoSQL • Schemaless; flexible/dynamic schema; Eventual Consistency; Strong Consistency; Tuneable consistency • Brewer's CAP Theorem • Sharding (partitioning) • Replication (Master/Slave; Peer-Peer) • Logical Ring, VNodes, Consistent Hashing, Hinted Hand-Off • Read-Repair • ACID Vs BASE • NoSQL types and common characteristics: Document Database(MongoDB); Key Value Store(Riak); Graph Databases; Column Family Store(Cassandra)
Warehousing Concepts	<ul style="list-style-type: none"> • OLTP systems VS. Data warehousing systems • A Typical Architecture (Inmon Vs. Kimball) OLTP, ODS, Detailed Data, Lightly and highly Summarised Data, ETL metadata repository, analysis tools etc. Summary Tables Vs Materialised Views • Data Warehouse Types-Enterprise DW; Data Mart, Virtual DW
Data Warehouse Design	<ul style="list-style-type: none"> • Logical Design <ul style="list-style-type: none"> ○ Star Schema <ul style="list-style-type: none"> ▪ Fact table makeup ▪ Dimension characteristics ○ Star Schema variants e.g. <ul style="list-style-type: none"> ▪ snowflake, starflake, fact constellations ○ Dimensions and multi-hierarchies; ○ Slowly changing dimension types; ○ Conformed dimensions – drill across • A Design Methodology to identify candidate star schemas from the corporate model <ul style="list-style-type: none"> ○ Classification Entities; Component Entities; Transaction Entities
Hadoop	<ul style="list-style-type: none"> • The 5 Vs of Big Data • Hadoop HDFS Storage: File Splitting/Slicing: Blocks; Rack Awareness Replication, Write operation Pipeline\ Read Operation NameNode, DataNode, SecondaryNameNode (2NN), Standby NameNode for HA, DataNodeFailure • Hadoop Processing: YARN, ResourceManager, Application Master Container, containers, NodeManager • MapReduce, MapReduce Tools e.g. HiveQL, PIG

Please see class notes for more details

Note: this is indicative list of what was covered and is not conclusive. Please refer to your class notes and Moodle notes for more details.