Name : .....

Roll No. : .....

Invigilator's Signature : .....

CS/M.TECH(CSE)/SEM-2/MCSE-201/2012

2012

ADVANCED DATABASE MANAGEMENT SYSTEM

Time Allotted : 3 Hours *Full Marks :* 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Attempt Q.No. 1 and *four* questions from the rest.

1. Attempt all questions : 5 × 2 = 10

   a) What is classification ? Differentiate between classification and clustering.

   b) What is association rule ? Give an example.

   c) Describe four properties of Data warehouse :

      (i) Subject-oriented

      (ii) Integrated

      (iii) Time variant

      (iv) Non-volatile.

   d) Suppose a group of 12 sales price records has been sorted as follows :

      5, 10, 11, 13, 15, 35, 50, 55, 72, 92, 204, 215

Partition them into three bins by each of the following methods :

(i) equal-frequency (equidepth) partitioning

(ii) equal-width partitioning

(iii) clustering.

e) (i) What do you mean by data replication and fragmentation of DDB ?

(ii) Explain different types of access methods in DDB.

2. a) What do you mean by smoothing by bin means and smoothing by bin medians ? 5

b) Compare OLAP and OLTP. 5

c) Explain the term 'Missing value', 'Noisy data' and 'Data cleaning'. 5

3. a) Write and explain three conditions for DDB fragmentation. What is predicate in DDB ? Generate minterm predicates considering p1 : CITY = "SF" and p2 : CITY = "LA". 4 + 1 + 2

b) Consider an airline reservation database :

(i) Flight description, include flight number, departure and arrival place, number of seats available.

(ii) Passenger description, including code, name, address, and phone number.

(iii) Reservation description, including passenger's code, flight number, and seat reserved. Distribute the database over two sites which are geographically located at New York, Atlanta. Fragmentation will be done either by arrival place or departure place.

Query : A request with flight number (IND2503) for flight availability .

(i) Write global quarry

(ii) Write fragmented quarry

(iii) Write allocation quarry. 8

4. a) Consider a distributed query like PJSUMSL Area = *N* or Area = *S* (SUPPLY JN DEPT) from tables SUPPLY (Supply_No, Product_No, Dept._No, Quantity, Dept (Dept_No, Dept_Name, Area).

Write down the Optimized query and show each steps for that. 7

b) Suppose you are given a relation *R* with four attributes ABCD.

(i) Identify the candidate key(s) for R.

(ii) Identify the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF)

(iii) If *R* is not in BCNF.

Decompose it into a set of BCNF relations that preserve the functional dependencies for each set for Functional dependence.

(I) $C \rightarrow D,\ C \rightarrow A,\ B \rightarrow C$

(II) $A \rightarrow B,\ BC \rightarrow D,\ A \rightarrow C$. 8

5. a) Write and explain *apriori* algorithm considering 9 transactions and 5 items and each transaction must have at least two items. 12

   b) Find out support() and confidence () at least rule. 3

6. WBUT wants to record the grades for the courses completed by students. There are four dimensions and consider two measured value (marks and remarks)

   (i) Course

   (ii) Student

   (iii) Professor

   (iv) Period.

   The only fact that is to be recorded in the table is course grade :

   a) Design star schema

   b) Using the above example describe the following OLAP operations :

   (i) Roll-up (ii) Slice. 10 + 5