School of Information Technology
IIT Kharagpur

Course Id: IT60107 Data Warehousing and Data Mining (Class Test 1)

Date: August 24, 2011
Total Time: 1 Hour   **Max. Marks: 20**

Answer all questions.

1. Consider a 3-D data array consisting of 3 dimensions A, B and C. The 3-D array is partitioned into 64 memory-based chunks. Dimension A is organized into 4-equisized partitions a0, a1, a2 and a3. Similarly dimensions B and C are also organized into 4-equisized partitions each. Chunks are numbered as 1, 2, 3, …, 64 corresponding to the sub cubes a0b0c0, a1b0c0, a2b0c0, a3b0c0, a0b1c0, …, a3b3c3, respectively. Suppose the sizes of the dimensions A, B and C are 2000, 3,000 and 1,000, respectively. If we perform multi-way array aggregation in the serial order 1….64 then calculate the minimum memory requirement for holding all relevant 2-D partial sums in memory.   **10 Marks**

2. Consider the following lattice of views along with a representation of the number of rows in each view where A is the base cuboid:

If you have to choose 3 views to materialize apart from the base cuboid, which of the views B-H would you choose and how? Assume that the cost of running a query is linearly proportional to the number of rows in the view from which it is run.   **10 Marks**