Class Test I - Repeat

Course Id: 406035 Data Warehousing and Data Mining

Date: September 17th, 2003.   Total Time: 60 minutes
Max. Marks: 60

An insurance company, with branches all over the country, wants to develop a data warehouse for effective decision-making about their insurance policies. There are a number of different types of insurance like Auto insurance, Home insurance, Industrial insurance, etc. The entire country is categorized into four regions, namely, North, South, East and West. Each region consists of a set of states. There may be different types of customers like individuals, institution, industry, etc. The data warehouse should record an entry for each policy issued to each customer along with the premium paid.

With respect to the above business scenario, answer the following questions. Clearly state any reasonable assumptions you make.

10+5+5+5+10+5+10+10

1. Design a star schema for the data warehouse clearly identifying the fact table(s), dimensional table(s), their attributes and measures along with the primary key and foreign key relationships.

2. Write an SQL query by which you can display region-wise, insurance-type-wise, year-wise total premium collected from your schema.

3. Draw a cuboid that would display the result of the query specified in Q. 2 above.

4. From the cuboid of Q. 3 above, if we want to see the amount of premium collected during the year 2001 for the state of Maharashtra for each type of customer, which sequence of OLAP operations would you need to perform?

5. Show the lattice of cuboids for the multi-dimensional data considering all the dimensions in your schema using a single level of hierarchy for each dimension.

6. Draw possible schema hierarchies for each dimension.

7. Based on the schema hierarchies drawn in Q. 6 above, determine the total number of cuboids, considering all the aggregation levels.

8. Once your data warehouse is ready and operational, there is a new requirement to maintain the amount of claim lodged at the same level of granularity. Extend your star schema to a fact constellation schema to take care of the new requirement.