

Answer ALL the Questions

1. Fill in the Blanks [1 x 10 =10]
 - a. Pre-computed _____ can solve performance problems
 - b. NUMA stands for _____
 - c. B-Tree is used as an index to provide access to records _____
 - d. When you _____ the data, you are aggregating the data to a higher level.
 - e. _____ are qualifying characteristics that provide additional perspectives to a given fact.
 - f. A _____ view is a dynamic table that contains not only the SQL query command to generate the rows, but also stores the actual rows.
 - g. _____ is the collection of data objects that are similar to one another within the same group
 - h. ETL stands for _____
 - i. _____ partitioning maximizes processing power and avoids bottlenecks
 - j. In _____ index, every key in the data file is represented in the index file.

2. Short Answers [2 x 2=4]
 - a. When Bit Map indexes are used?
 - b. How do data warehouses achieve fast performance?

3. Long Answers
 - a. Which one is faster, Multidimensional OLAP or Relational OLAP and why?
[2 + 2= 4]

 - b. Design the data warehouse for a wholesale furniture company. The data warehouse has to allow to analyze the company's situation at least with respect to the Furniture, Customers and Time. Moreover, the company needs to analyze: the furniture with respect to its type (chair, table, wardrobe, cabinet. . .), category (kitchen, living room, bedroom, bathroom, office. . .) and material (wood, marble. . .), the customers with respect to their spatial location, by considering at least cities, regions and states The company is interested in learning at least the quantity, income and discount of its sales [10 + 6 = 16]

- a) Draw a star schema for the data warehouse
- b) SQL queries:
 - (1) Find the quantity, the total income and discount with respect to each city, type of furniture and the month
 - (2) Find the average quantity, income and discount with respect to each country, furniture material and year
 - (3) Determine the 5 most sold furnitures during the May month
- c. Discuss different optimization techniques applied to data Warehouse with respect to [3 x 2 = 6]
 - a) Query
 - b) Hardware resource
 - c) Network resource
- d. Suppose that a state's tourism department wants to improve its revenue by promoting different tourism packages associated with different tourist places. But the department needs a warehouse which can keep track of such information, amount spent towards such activities, revenue generated and hence the profit. The Warehouse needs the various dimensions and fact to achieve this. If you were given the task of building such a warehouse how would you choose attributes & measures for building the schema [6+4=10]
 - a) Build a snowflake schema
 - b) Why metadata is important? Provide metadata for the above warehouse
- e. We have, [4]
001:10,Smith,Joe,40000; 002:12,Jones,Mary,50000;
003:11,Johnson,Cathy,44000; 004:22,Jones,Bob,55000;
Convert this into Column based