**Subject: Data Warehousing & Mining**

**Chapter Wise Questions**

| Module I | Information requirements are recorded for 'Hotel occupancy' considering dimensions like Hotel, Room and Time. Few Facts recorded are vacant room, occupied rooms, number pf occupants etc.<br><br>Answer the following questions for this problem:<br><br>i. Design the star schema<br><br>ii. Can you convert the star schema to snowflake schema? If yes, justify and draw the snowflake schema. | [10]<br><br>Nov 2018 |
|---|---|---|
| | College Wants to record the Marks for the courses completed by students using the dimensions:  i) Course ii) Student iii) Time & a measure Aggregate marks.<br><br>Create a Cube and perform following OLAP operations.<br><br>i)Rollup  ii) Drill down  iii) Slice  iv) Dice   v) Pivot | [10]<br><br>Nov 2018<br><br>May 17<br><br>May 2018 |
| | Discuss data warehousing design strategies in detail | [10]<br><br>Nov 2018<br><br>May 2018 |
| | Discuss various OLAP models and their architecture | [10]<br><br>Nov 2018 |
| | Updates to dimension tables | [5]<br><br>Nov 2018 |
| | Consider following dimensions for a supermarket chain: Product, Store, Time and Promotion. With respect to this business scenario, answer the following questions. Clearly state any reasonable assumptions you make. | [10]<br><br>Dec 2019 |

Parshvanath Charitable Trust's

**A. P. SHAH INSTITUTE OF TECHNOLOGY**
(Approved by AICTE New Delhi & Govt. of Maharashtra, Affiliated to University of Mumbai)
(Religious Jain Minority)

| | |
|---|---|
| a) Design an information package diagram for this business scenario.<br><br>b) Design a snowflake schema for the data warehouse, clearly depict the fact table(s). Dimension tables their attributes and measures. | |
| Consider a Data ware house for a sport manufacturing company storing sales details of various sports equipments sold and the time of the sale. Using this example describe the following OLAP operations:<br><br>i)Rollup ii) Drill down iii) Slice iv) Dice v) Pivot | [10]<br><br>Dec 2019 |
| Differentiate between top-down and bottom-up approaches for building data warehouse. Discuss the merits and limitations of each approach. Also explain the practical approach for designing a data warehouse. | [10]<br><br>Dec 2019 |
| Write short note on Metadata in Data Warehouse | [5]<br><br>Desc 2019<br><br>May 2017 |
| Consider following dimensions for a supermarket chain: Product, Store, Time and Promotion. With respect to this business scenario, answer the following questions. Clearly state any reasonable assumptions you make.<br><br>a) Design the star schema<br><br>b) Can you convert the star schema to snowflake schema? If yes, justify and draw the snowflake schema. Clearly depict the fact table(s). Dimension tables their attributes and measures. | [10]<br><br>Dec 2016 |
| What is meant by metadata in the context of a Data Warehouse? Explain the different types of meta data stored in a data. Illustrate with example, | [10]<br><br>Dec 2016 |
| Discuss how computations can be performed efficiently on data cubes | [10]<br><br>Dec 2016 |

| | | |
|---|---|---|
| | **Differentiate** <br><br> OLTP vs OLAP <br><br> Data Warehouse vs Data Mart | [10] <br><br> Dec 2016 <br><br> May 17 |
| Module I | Write short notes on <br><br> 1. Fact less Fact Table  2. Indexing OLAP data | [5] each <br><br> Dec 2016 |
| | What is dimension modelling? Design the data warehouse for Wholesale furniture Company. The data warehouse has to allow analyzing the company situation at least with respect to the Furniture , Customer and Time. More ever the company needs to anlayse  the furniture with respect to its  type, category and material. The customers with respect to thrie spatial location by considering at least cities regions and states. The company is interested in learning the quantity, income and discount of its sales. | [10] <br><br> May 17 |
| | Architecture of Data warehouse and differentiate Data warehouse and Data Mart | [10] <br><br> May 2018 <br><br> May 17 |
| | Information requirements are recorded for 'Hotel occupancy' considering dimensions like Hotel, Room and Time. Few Facts recorded are vacant room, occupied rooms, number pf occupants etc. <br><br> Answer the following questions for this problem: <br><br> i.      Design the star schema and snowflake schema <br><br> ii.     Calculate the maximum number of base fact table records for the values given <br><br> Time period 5 yaers <br><br> Hotels 150 <br><br> Rooms :  750 rooms in each hotel (About 400 occupied in each hotel daily) | [10] <br><br> May 2018 |

| Module I | What is Metadata? Why do we need metadata when search engines like Google seem so effective? | |
|---|---|---|
| | Suppose that a data warehouse consists of the three dimensions time, doctor and patient, and the two measures count and charge, where charge is the fee that a doctor charges a patient for a visit.<br><br>Draw a star schema diagram for the above data warehouse.<br><br>Starting with the base  cuboid [day doctor patient ] what specific OLAP operation should be performed in order to list the total fee collected by each doctor in 2010.<br><br>To obtain the same list write an SQL query assuming the data are stored in a relational database with the schema fee ( day month year doctor hospital patient count charge) | [10] |
| | Why is entity-relationship modeling technique not suitable for the data warehouse? How is dimensional modeling different? | [10] |