

B.Sc. (I. T.) Sem. - V (CBCS - 2015 Course) : SUMMER - 2019
SUBJECT: DATA WAREHOUSING AND DATA MINING

Day: Thursday
Date: 16/05/2019

S-2019-1285

Time: 02.30 p.m. to 05.30 p.m.
Max. Marks: 60

N.B.:

- 1) Attempt **ANY SIX** questions.
- 2) Figures to the right indicate **FULL** marks.
- 3) Draw appropriate figures **WHEREVER** necessary.

-
- Q.1** a) Explain fact table and dimension table with an example. (04)
b) Explain different types of schema with an example. (06)
- Q.2** a) What are the various design and implementation issues of OLTP? (04)
b) Explain materialized view with an example. (06)
- Q.3** a) Explain concept of Data, Information and Knowledge in KDD process with suitable example. (06)
b) With respect to Association Rule Mining define: (04)
i) Support
ii) Confidence
- Q.4** What is Dimensional modeling? Design the data warehouse for wholesale furniture company. The data warehouse has to allow analyzing the company's situation at least with respect to the Furniture, Customer and Time. Moreover, the company needs to analyze: (10)
i) The furniture with respect to its type, category and material.
ii) The customers with respect to their spatial location, by considering at least cities, regions and states. The company is interested in learning the quantity, income and discount of its sales.
- Q.5** Suppose that the data for analysis includes the attribute age. The values for attribute age for the data tuples are 4,8,15,21,24,25,28,34. Using the following binning methods for data smoothing, show the resultant data. (10)
i) Bin by medians
ii) Bin by boundaries
iii) Bin by means
Illustrate your steps.
- Q.6** 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. (10)
i) Enumerate three classes of schemas that are popularly used for modeling data warehouses.
ii) Draw a schema diagram for the above data warehouse using one of the schemas classes listed in (i).
iii) Starting with the base cuboid [day; doctor; patient], what specific OLAP operations should be performed in order to list the total fee collected by each doctor in 2004?
iv) To obtain the same list, write an SQL query assuming the data is stored in a relational database with the schema fee (day, month, year, doctor, hospital, patient, count, charge).
- Q.7** Explain Logical view Vs Materialized view with examples. What is the necessity of creating materialized view? Mention its advantages and disadvantages. (10)
- Q.8** Write short notes on: (10)
a) Star Schema
b) Data Mart