

Total No. of Questions : 12]

SEAT No. :

P1447

[4759]-203

[Total No. of Pages : 3

B.E.IT

BUSINESS INTELLIGENCE

(2008Course) (Elective-IV) (Semester-II)(414451)

Time : 3 Hours]

[Max. Marks : 100

Instructions to the candidates:

- 1) *Answers to the two sections should be written in separate answer books.*
- 2) *Answer any three questions from each section .*
- 3) *Neat diagrams must be drawn wherever necessary.*
- 4) *Figures to the right side indicate full marks.*

SECTION-I

- Q1)** a) Explain Structured and unstructured data. What are the challenges in implementation of Data Warehouse?. [8]
- b) Explain Subject Orientation and Application Orientation in data warehouse. Give example. [8]

OR

- Q2)** a) Explain usage of data warehouse in retail sales system. [8]
- b) How is an OLTP system useful in an organization? What are its applications. [8]

- Q3)** a) DW has one dimension called 'Product' with attributes like product_ID, Product_Name, Product_Category.
The product ice cream has Product_ID =15, Product_Name= "Ice. Cream" and Product_Category= "Dairy".

On 1st Apr 2014 the product category of ice cream is changed to

"Cold Storage" Elaborate 3 ways to handle this scenario [8]

- b) Write a short notes on
- i) Star Schema
 - ii) Snowflake Schema.

[8]

OR

P.T.O.

- Q4)** a) List out the important characteristics of a dimension table.Explain conformed dimensions. [8]
- b) Explain transaction and recurring snapshot types of dimensional modeling. Highlight two main differences between these two modeling types. [8]
- Q5)** a) What is ETL? Explain the architecture of ETL. [9]
- b) Explain the following terms [9]
- i) Data Extraction & Cleansing.
 - ii) Data scrubbing
 - iii) Cubes

OR

- Q6)** a) What is data Staging? Explain its Explain Its pros and cons. [6]
- b) How are following scenarios handled in the ETL process: [6]
- i) NULL values are present in the Operation Data Source.
 - ii) Multiple abbreviations for the same value e,g. Pune is represented as PUN, POONA, Pune etc in Operational Data Source.
- c) Data sources form which the data is extracted for ETL process can be heterogeneous at many times. What are the typical challenges involved in handling such data sources? [6]

SECTION II

- Q7)** a) What is OLAP? Explain “slicing and dicing”, “drill down and Roll-up” and “Multidimensional” with respect to the OLAP. [9]
- b) Explain reporting architecture with suitable diagram. [7]

OR

- Q8)** a) What is the importance of visualization in BI system? Explain any 4 types of charts including their purpose and significance in details. [8]
- b) Explain different levels of securities implemented in DW applications.Explain report level security in details. [8]

- Q9)** a) List different statistical techniques for data analysis. Explain any one of them. [6]
- b) Explain any one method of hierarchical clustering with an example. [10]

OR

- Q10)** a) There are 2 sets of data mining algorithms - supervised and unsupervised. Elaborate the major difference between the 2. Clustering is one of them. Is it supervised or unsupervised? Why? Explain any one technique that is used for clustering numeric data. [10]
- b) Given the data from the stock market it needs to be predicted that how a particular stock will perform in future. Which analysis technique can be used? Justify. [6]

- Q11)** a) What is BIG Data? List different technologies available for supporting BIG data. [9]
- b) Explain following terms
- i) Operational BI.
 - ii) Agile BI. [9]

OR

- Q12)** a) Explain BI on Cloud. [2]
- b) Explain real time BI. Explain its application in detail. [9]
- c) Explain salient features of Netezza. [7]

