

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code : 31284

B.E./B.Tech. DEGREE EXAMINATION, NOVEMBER/DECEMBER 2013.

Seventh Semester

Computer Science and Engineering

CS 2032/CS 701/10144 CSE 32 — DATA WAREHOUSING AND DATA MINING

(Common to Sixth Semester Information Technology)

(Regulation 2008/2010)

Time : Three hours

Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. Differentiate between MOLAP and ROLAP.
2. List any two tools for performing OLAP.
3. What is metadata?
4. What is ETL process? Give its significance.
5. Differentiate between data characterization and discrimination.
6. Give the need for data pre-processing.
7. Define frequent itemset.
8. Give examples for binary and multidimensional association rules.
9. How the goodness of clusters is measured?
10. Highlight few applications involving outlier analysis.

PART B — (5 × 16 = 80 marks)

11. (a) Draw any two multi-dimensional schemas suitable for representing weather data and give their advantages and disadvantages. (16)

Or

- (b) Explain the multi-tier architecture suitable for evolving a data warehouse with suitable diagram. (16)

12. (a) Write a short note on the following :
- (i) Features of Cognos tool. (8)
 - (ii) Multidimensional data model. (8)

Or

- (b) Discuss how data warehousing is used in retail and telecommunication industry. (16)

13. (a) Write short notes on :
- (i) Classification of data mining systems (8)
 - (ii) Data mining task primitives. (8)

Or

- (b) Write short notes on the various pre-processing tasks. (16)

14. (a) Discuss about mining association rules using the apriori algorithms. (16)

Or

- (b) Explain as to how neural networks are used for classification of data. (16)

15. (a) Explain any two hierarchical clustering techniques stating their pros and cons. (16)

Or

- (b) Write short notes on :
- (i) DBSCAN (6)
 - (ii) Outlier analysis. (10)