

N. B. : (1) Question No.1 is compulsory.

(2) Answer any **four** out of the remaining questions.

(3) Answer to **sub** questions must be **written** together.

1. (a) What are the different characteristics of a Data Warehouse? 5
- (b) For a Supermarket Chain consider the following dimensions, namely 5
Product, store, time, promotion. The schema contains a central fact table, sales facts with three measures unit_sales, dollars_sales and dollar_cost. Design star schema for this application.
- (c) Explain Web usage mining. 5
- (d) Illustrate how the supermarket can use clustering methods to improve sales. 5
2. Define the following terms :- 20
 - (a) Dimension Tables
 - (b) Snowflake Schema
 - (c) Web Structure Mining
 - (d) Supervised learning
3. (a) Explain Hierarchical Clustering methods. 10
- (b) Explain the Page Rank algorithm. 10
4. (a) Describe the following OLAP operations using an example; 10
 - (1) Slice
 - (2) Dice
 - (3) Rollup
 - (4) Drill Down
 - (5) Pivot
- (b) Consider the following transaction database: 10

TID	Items
01	A,B,C,D
02	A,B,C,D,E,G
03	A,C,G,H,K
04	B,C,D,E,K
05	D,E,F,H,L
06	A,B,C,D,L
07	B,I,E,K,L
08	A,B,D,E,K
09	A,E,F,H,L
10	B,C,D,F

[TURN OVER

Apply the Apriori algorithm with minimum support of 30% and minimum confidence of 70% and find all the association rule in the data set. 10

5. (a) Explain Classification Algorithms 10
(b) Explain the ETL (Extract, Transform Load) cycle. 10
6. (a) Define multidimensional and multilevel association mining. 10
(b) Explain the role of Meta data in a data warehouse. 10
7. (a) Write detailed notes on 20
(a) Data Warehouse Architecture
(b) K-Means Clustering