

No. of Printed Pages : 4

Sl. No.

B3.E4-R5 : DATA WAREHOUSING AND DATA MINING

DURATION : 03 Hours

MAXIMUM MARKS : 100

Roll No. :

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

Answer Sheet No. :

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

Name of Candidate : _____ ; **Signature of Candidate :** _____

INSTRUCTIONS FOR CANDIDATES :

- Carefully read the instructions given on Question Paper, Answer Sheet.
- Question Paper is in English language. Candidate has to answer in English Language only.
- Question paper contains Seven questions. The Question No. 1 is compulsory. Attempt any FOUR Questions from Question No. 2 to 7.
- Parts of the same question should be answered together and in the same sequence.
- Questions are to be answered in the ANSWER SHEET only, supplied with the Question Paper.
- Candidate cannot leave the examination hall/ room without signing on the attendance sheet and handing over his/her Answer Sheet to the Invigilator. Failing in doing so, will amount to disqualification of Candidate in this Module/Paper.
- After receiving the instruction to open the booklet and before answering the questions, the candidate should ensure that the Question Booklet is complete in all respects.

DO NOT OPEN THE QUESTION BOOKLET UNTIL YOU ARE TOLD TO DO SO.

1. (a) What is Data Mining ? How does it relate to Data Warehousing ?
- (b) What are the major issues in Data Warehousing ?
- (c) What do you mean by Partitioned clustering ?
- (d) What is Web usage mining ?
- (e) Discuss in short about Data warehouse architecture.
- (f) What is outlier analysis ? Why is outlier mining important ?
- (g) Calculate the two regression equations of X on Y from the data given below, taking deviations from the actual means of X and Y.

Price (Rs.)	10	12	13	12	16	15
Amount demanded	40	38	43	45	37	43

(7x4)

2. (a) Elaborate the term Data Warehouse and why it is important ?
 - (b) Discuss data model and its types in detail.
 - (c) What is the purpose of data mart ? Explain.
- (8+6+4)
3. (a) What is ETL in data warehouse ? Explain the types of ETL tools in detail.
 - (b) Differentiate between ETL and ELT processes.
 - (c) What is Change Data Capture Technique ? Explain its benefits.
- (10+4+4)
4. (a) Why data cleaning is necessary ? Discuss with the help of example.
 - (b) What is Data Mining Query Language ? Explain in detail.
 - (c) Explain KDD process in detail.
- (6+6+6)
5. (a) Compare MOLAP Model and ROLAP Model.
 - (b) Differentiate between OnLine Transactional Processing (OLTP) and OnLine Analytical Processing (OLAP).
 - (c) What is the relationship between Data Warehousing and Data Mining ?
- (8+4+6)

6. (a) Explain decision tree classification with the help of suitable example.
(b) Explain k nearest neighbor classifier with an illustration.
(c) What is Nonlinear Regression ? Give its expression. (6+8+4)
7. (a) Explain the Apriori Algorithm and its role in Association Rule Mining.
(b) What is Hierarchical Clustering ? Explain in detail.
(c) Elaborate the steps involved in data preprocessing. (6+8+4)

- o O o -

SPACE FOR ROUGH WORK