



Course/Branch : CS (AIDS)/AI & ML/AI&DS Semester : 7<sup>th</sup>  
 Subject Name : Data Warehousing and Data Mining Max. Marks : 100  
 Subject Code : KA1075 Time : 180min

**CO-1** : Be familiar with mathematical foundations of data mining tools.

**CO-2** : Understand and implement classical models and algorithms in data warehouses and data mining

**CO-3** : Characterize the kinds of patterns that can be discovered by association rule mining, classification and clustering.

**CO-4** : Master data mining techniques in various applications like social, scientific and environmental

context.

**CO-5** : Develop skill in selecting the appropriate data mining algorithm for solving practical problems.

**Section – A # 20 Marks (Short Answer Type Questions)**

Attempt ALL the questions. Each Question is of 2 marks (10 x 2 = 20 marks)

Q. No.	COx	Question Description # Attempt ALL the questions. Each Question is of 2 marks
1	a	CO1 Explain metadata in the context of a data warehouse. (BKL : K2).
b	CO1	CO1 Mention one difference between a database system and a data warehouse. (BKL : K1).
c	CO2	CO2 Demonstrate a data warehousing strategy. (BKL : K3)
d	CO2	CO2 Explain client/server computing model in data warehousing. (BKL : K3)
e	CO3	CO3 Name any two functionalities of data mining. (BKL : K1)
f	CO3	CO3 Explain the purpose of binning in data cleaning? (BKL : K2)
g	CO4	CO4 Define data generalization in classification. (BKL : K3)
h	CO4	CO4 Write the main goal of association rules in data mining. (BKL : K3)
i	CO5	CO5 Name the three types of OLAP servers. (BKL : K1).
j	CO5	CO5 Mention one use of historical information in data warehousing. (BKL : K1).

**Section – B # 30 Marks (Long / Medium Answer Type Questions)**

Attempt ALL the questions. Each Question is of 6 marks (5 x 6 = 30 marks)

Q.2 (CO-1) : Describe the steps involved in building a data warehouse in detail. (K2)

OR

Compare a star schema differ from a snowflake schema in terms of structure and performance. (K2)

Q.3 (CO-2) : Explain the role of operating systems in a data warehouse environment. (K3)

OR

Write the main steps involved in planning and implementing a data warehouse. (K3)

Q.4 (CO-3) : Elaborate how does regression help in cleaning noisy data? Provide an example (K2)

OR

Explain the concept of data cube aggregation and its use in data reduction. (K2)

Q.5 (CO-4) : Describe the working of decision tree-based algorithms for classification with an example. (K3)

OR

Explain the STING method in grid-based clustering and its applications. (K3)

Q.6 (CO-5) : Explain the key differences between ROLAP, MOLAP, and HOLAP (K2)

OR

Describe the role of backup and recovery in maintaining data warehouse reliability. (K2)

**Section – C # 50 Marks (Medium / Long Answer Type Questions)**

Attempt ALL the questions. Each Question is of 10 marks.

Q.7 (CO-1) : Attempt any ONE question. Each question is of 10 marks.

- Describe in detail the components of a data warehouse and their roles in the overall architecture. (K2)
- Compare and contrast star schema, snowflake schema, and fact constellation schema in terms of design complexity and query performance. (K2)

Q.8 (CO-2) : Attempt any ONE question. Each question is of 10 marks.

- Write the hardware requirements for implementing a robust data warehouse and their role in optimizing performance. (K3)
- Explain in detail about distributed DBMS implementations and cluster systems in data warehousing. (K3)

Q.9 (CO-3) : Attempt any ONE question. Each question is of 10 marks.

- Compare and contrast various data preprocessing techniques (e.g., data cleaning, dimensionality reduction, and data transformation) and their impact on data mining outcomes. (K2)
- Write a detailed note on data reduction techniques, including data compression and numerosity reduction. (K3)

Q.10 (CO-4) : Attempt any ONE question. Each question is of 10 marks.

- Write a detailed note on the statistical-based and distance-based algorithms for classification, highlighting their use cases. (K3)
- Explain the neural network approach in association rule mining and its advantages over traditional methods. (K3)

Q.11 (CO-5) : Attempt any ONE question. Each question is of 10 marks.

- Write a detailed note on OLAP functions and tools, including examples of their applications. (K2)
- Compare and contrast spatial mining and temporal mining, highlighting their real-world applications (K2)