MASTER OF COMPUTER APPLICATIONS (MCA_NEW/MCAOL)

SEMESTER-2 ASSIGNMENTS OF MCA_NEW/MCAOL PROGRAMME S

(July - 2026 & January - 2027)

MCS-218, MCS-219, MCS-220, MCS-221

MCSL-222, MCSL-223

SCHOOL OF COMPUTER AND INFORMATION SCIENCES
INDIRA GANDHI NATIONAL OPEN UNIVERSITY
MAIDAN GARHI, NEW DELHI – 110 068

CONTENTS

Important Notes

1. MCA_NEW Student need to submit the assignments to the Coordinator of your Study Centre on or before the due date.
2. MCAOL Student need to submit the assignments on LMS Portal on or before the due date.
3. Assignment submission before due dates is compulsory to become eligible for appearing in corresponding Term End Examinations. For further details, please refer to Programme Guide.
4. To become eligible for appearing the Term End Practical Examination for the lab courses, it is essential to fulfill the minimum attendance requirements as well as submission of assignments (on or before the due date). For further details, please refer to the Programme Guide.
5. The viva voce is compulsory for the assignments. For any course, if a student submitted the assignment and not attended the viva-voce, then the assignment is treated as not successfully completed and would be marked as ZERO.

Course Code	**:**	**MCS-221**
Course Title	**:**	**Data Warehousing and Data Mining**
Assignment Number	**:**	**MCA_NEW/MCAOL(II)/221/Assign/2026-27**
Maximum Marks	**:**	**100**
Weightage	**:**	**30%**
Last Date of Submission	**:**	**31st October, 2026 (For July, 2026 Session)** **15th April, 2027 (For January, 2027 Session)**

Note: This assignment has ten questions. All the questions are compulsory and there is no choice. Rest 20 marks are for viva voce. You may use illustrations and diagrams to enhance the explanations. Please go through the guidelines regarding assignments given in the Programme Guide.

Q1: A university intends to build an enterprise data warehouse by integrating admissions, examinations, finance and LMS databases. Explain the need for a data warehouse, discuss its characteristics, and justify whether the Inmon or Kimball approach is more suitable. **(8 Marks)**

Q2: Illustrate a three-tier data warehouse architecture for a healthcare organization. Explain the role of data sources, staging area, ETL, metadata repository, warehouse, OLAP server and front-end tools. **(8 Marks)**

Q3: A supermarket chain wants to analyse monthly sales, customer behaviour and product performance. Design a dimensional model identifying fact table(s), dimensions, measures and appropriate schema. Justify your design. **(8 Marks)**

Q4: Explain the complete ETL process for integrating data from e-commerce, ERP and CRM systems into a data warehouse. Discuss common data quality issues and suitable transformation strategies. **(8 Marks)**

Q5: Compare OLTP and OLAP systems with suitable examples. Discuss OLAP operations (roll-up, drill-down, slice, dice and pivot) and explain how they support managerial decision making. **(8 Marks)**

Q6: A financial institution has noisy, missing and inconsistent customer data. Explain suitable data preprocessing techniques including cleaning, integration, transformation, reduction and discretization before mining. **(8Marks)**

Q7: Explain the working of association rule mining using an illustrative transaction dataset. Describe the Apriori algorithm and interpret support, confidence and lift for generated rules. **(8Marks)**

Q8: Compare decision tree, Naïve Bayes, k-nearest neighbour and support vector machine classifiers. Recommend the most appropriate algorithm for credit-risk prediction with justification. **(8 Marks)**

Q9: Differentiate partitioning, hierarchical and density-based clustering techniques. Discuss the working of K-means and DBSCAN and compare their suitability for real-world applications involving noisy data. **(8 Marks)**

Q10: Discuss emerging trends in data warehousing and data mining such as cloud data warehouses, big data analytics, text mining, web mining and data stream mining. Explain how these technologies enhance business intelligence. **(8 Marks)**