M.P. Birla Institute of Management LIBRARY

62239

Reg. No. | | | | | | | | |

III Semester M.B.A. (Day / Evening) Degree Examination, May/June - 2025

MANAGEMENT

Data Warehousing and Data Mining

(CBCS Scheme 2019 Onwards)

Paper : 3.7.3

Time : 3 Hours **Maximum Marks : 70**

SECTION - A

Answer any **Five** questions from the following each carries **5** marks. **(5×5=25)**

1. Compare the benefits of employing cloud services for Data Warehousing versus Traditional on - premise options.
2. Determine the significance of effective team organisation and role definition throughout the data warehouse development life cycle.
3. Examine the impact of improper Data transformation methods on the overall accuracy of a Data Warehouse.
4. Evaluate the design of a Data Mining system and explain how each component helps with knowledge discovery.
5. Explore the application of Data Mining in the healthcare industry for early disease detection and patient profiling.
6. Analyse the primary distinctions between data lakes and data warehouses in terms of data structure, use cases and performance.
7. Using a sales dataset, demonstrate how to manage missing data and delete duplicates in an ETL process.

SECTION - B

Answer any **Three** questions from the following each carries **10** marks. **(3×10=30)**

8. Assess the usefulness of data marts in addressing specific business unit requirements in a data warehousing context.
9. Analyse the distinctions between MOLAP, ROLAP, and HOLAP designs and determine their applicability for various sorts of data access requirements.

[P.T.O.

CamScanner

10. Examine the capability of a popular ETL solution (e.g., Talend, informatica, Apache NiFi) for automating the ETL procedure for large - scale corporate data sets.

11. Evaluate the influence of Data mining in retail, with a focus on how consumer segmentation enhances focused marketing techniques.

SECTION - C

12. Compulsory Case Study : **(1×15=15)**

NovaBank, a regional bank, was experiencing issues with its legacy on - premise data warehouse systems. These systems were expensive to maintain, rigid, and delayed in providing insights for real - time decision - making. NovaBank modernised by using a cloud - based managed data warehouse service, Amazon Redshift, to move their existing data architecture.

To accommodate future scalability and different data sources, they created a data lake architecture on AWS S3, allowing them to store structured and unstructured data at scale. The new environment allowed for faster querying, improve analytics, and greater integration with BI tools. NovaBank may now merge consumer transaction, digital engagement, and third - party data to create a 360- degree customer view.

However, the transition was not without issues. Data governance issues, privacy rules (such as GDPR), and the necessity for experienced cloud specialists all presented substantial obstacles. In addition, the bank had to train its internal personnel and update access control measures to ensure data security and compliance.

Questions :

a. Use your knowledge to determine how the development of a data lake and a cloud data warehouse aided NovaBank's data strategy.

b. Analyse the primary risks that NovaBank encountered throughout its shift to a cloud - based architecture. What solutions could they implement to address these difficulties?

c. Assess the efficiency of managed services like Amazon Redshift in supporting Digital Transformation projects in conventional industries such as banking.

CamScanner