

Total No. of Questions : 8]

SEAT No. :

**PB2511**

[Total No. of Pages : 3

[6263]-397

**B.E. (Artificial Intelligence and Data Science)  
BUSINESS INTELLIGENCE  
(2019 Pattern) (Semester-VIII) (417533 B) (Elective-VI)**

*Time : 2½ Hours]*

*[Max. Marks : 70*

*Instructions to the candidates:*

- 1) Answer Q.1 or Q.2, Q.3 or Q.4, Q. 5 or Q. 6, Q. 7 or Q. 8.
- 2) Neat diagrams must be drawn wherever necessary
- 3) Assume suitable data if necessary.
- 4) Figures to the right indicate full marks.

**Q1)** a) Evaluate the impact of Data Duplication on the overall performance and efficiency of a data warehouse [8]

- b) Explain the concepts of ETL Architecture, Extraction, Transformation, and Loading (ETL) in the context of Business Intelligence. [5]
- c) Explain the concept of Change Data Capture and its significance in data provisioning. [5]

OR

**Q2)** a) Analyse the role of Lookups in the transformation process of ETL. [8]

b) Create a simplified data mart and explain its purpose in a BI system. [5]

c) Describe the key components of Data Provisioning, including Data Quality, Data Profiling, and Data Enrichment. [5]

**Q3)** a) Evaluate the effectiveness of different methods of Data discretization in improving the interpretability of data. [8]

b) Describe the significance of Data Reduction in the context of large datasets, and explain the methods of Sampling. [5]

c) Identify the types of data affected by noise and explain their impact on data quality. [4]

OR

**P.T.O.**

- Q4)** a) Compare and contrast Feature Selection and Principal component Analysis (PCA) in terms of their applicability and limitations. [8]  
 b) Differentiate between univariate, Bivariate and Multivariate analysis. [5]  
 c) Discuss the importance of Data Exploration in the data analysis pipeline. [4]

- Q5)** a) Consider the following dataset and we will find frequent item sets and generate association rules for them using apriori algorithm, consider minimum support count is 2 & minimum confidence is 60% [9]

TID	Items
T1	11, 12, 15
T2	12, 14
T3	12, 13
T4	11, 12, 14
T5	11, 13
T6	12, 13
T7	11, 13
T8	11, 12, 13, 15
T9	11, 12, 13

- b) Define Bayesian methods and Logistic Regression in the context of classification. [5]  
 c) List and explain the key components involved in the evaluation of regression models. [4]

OR

- Q6)** a) Explain logistic regression with example considering relevant variables and data. [9]  
 b) Discuss the principles underlying the Apriori Algorithm for association rule mining. [5]  
 c) Define Hierarchical methods in clustering and describe their characteristics. [4]

- Q7)** a) Evaluate the benefits and challenges of integrating Cloud Computing with Business Intelligence. [9]  
b) Explain the importance of Business Intelligence in enhancing Customer Relationship Management strategies. [5]  
c) Provide examples of Business Intelligence applications in the Manufacturing Industry. [3]

OR

- Q8)** a) Create a plan for integrating Web 2.0 and Online Social Networking features into existing Business Intelligence systems. [9]  
b) Describe the impact of Business Intelligence on Healthcare Monitoring processes and outcomes. [5]  
c) List and briefly explain the applications of Business Intelligence in the Higher Education sector. [3]