17299

Library
Brainware University
398, Ramkrishnapur Road, Barasat
Kolkata, West Bengal-700125

BRAINWARE UNIVERSITY

Term End Examination 2025-2026
Programme – B.Tech.(CSE)-DS-2021/B.Tech.(CSE)-DS-2022/B.Tech.(CSE)-DS-2023
Course Name – Data Analytics and Reporting
Course Code - PCC-CSD503
(Semester V)

Full Marks : 60 **Time : 2:30 Hours**

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A
(Multiple Choice Type Question) 1 x 15=15

1. *Choose the correct alternative from the following :*

(i) Identify the name of the first phase of data analytics lifecycle.
a) Model planning b) Data preparation
c) Discovery/objective d) Data visualization

(ii) Select the name of the tool that is NOT used for data preparation:
a) Power BI b) Hadoop
c) Python d) Tableau

(iii) Choose the name of the phase that the analyst team delivers final report, briefings and codes .
a) Operationalize b) Model Building
c) Model Planning d) None of these

(iv) Identify the name that is not the dimension of Data Quality.
a) Accuracy b) Completeness
c) Validity d) Volume

(v) Select the correct example of Ordinal data.
a) Price b) Temperature
c) Education Level d) All of these

(vi) Select the correct property from the following option for nominal data.
a) Order b) Grouping
c) Equal interval d) None of these

(vii) Choose the proper BI tool that enable users to analyse data along multiple dimensions.
a) Ad hoc b) OLAP
c) Tableau d) None of these.

(viii) Choose the correct option regarding the benefits of Business Intelligence.
a) Speeds up and improves decision-making b) Optimizes internal business processes
c) Increases operational efficiency and productivity d) All of these

248- 4

(ix) Select the primary purpose of a histogram in data analysis.

a) To display individual data points
b) To show the relationship between two variables
c) To summarize the distribution of a dataset
d) To calculate the mean and median of the data

(x) Choose the main advantage of using a histogram over a bar chart:

a) A histogram can display multiple variables simultaneously while a bar chart cannot.
b) A histogram shows the distribution of continuous data while a bar chart compares discrete data categories.
c) A histogram is better at showing relationships between variables compared to a bar chart.
d) A histogram is more suitable than a bar chart for displaying categorical data.

Library
Brainware University
398, Ramkrishnapur Road, Barasat
Kolkata, West Bengal-700125

(xi) Choose the primary goal of data analytics.

a) To collect data
b) To store data
c) To process data
d) To gain insights from the data

(xii) Choose the data pre-processing technique in data analytics.

a) Data visualization
b) Data cleaning
c) Data modeling
d) Data reporting

(xiii) Choose the type of data analysis involved in discovering patterns or relationships in data to make predictions about future events:

a) Descriptive analytics
b) Diagnostic analytics
c) Predictive analytics
d) Prescriptive analytics

(xiv) Choose the purpose of Exploratory Data Analysis (EDA).

a) To summarize data
b) To make predictions
c) To find hidden patterns
d) To create reports

(xv) Choose the programming language commonly used for data analysis and manipulation.

a) Java
b) Python
c) C++
d) Ruby

Group-B

(Short Answer Type Questions) 3 x 5=15

2. Discuss the characteristics of Big Data. (3)
3. Illustrate MapReduce in the context of Hadoop, and the method of its process data. (3)
4. Explain the advantages of Hadoop. (3)
5. Define data warehouse, and its relation to Business Intelligence. (3)
6. Analyze the impact of symbol generalization in thematic mapping. (3)

OR

Differentiate between qualitative and quantitative thematic maps with examples of each usage. (3)

Group-C

(Long Answer Type Questions) 5 x 6=30

7. Discuss Big Data and its types. (5)
8. Describe key advantages of using Hadoop and MapReduce for big data processing (5)
9. Discuss about the purpose of mapper and reducer in MapReduce component in Hadoop. (5)
10. Explain Hadoop and evaluate its elements. (5)
11. Express the advantages and limitations of unsupervised learning algorithms in data analysis. (5)
12. Justify the importance of the modeling process in machine learning. (5)

OR

Express the significance of validating a machine learning model before deployment. (5)