



BRAINWARE UNIVERSITY

Term End Examination 2025-2026

Programme – B.Tech.(CSE)-DS-2022/B.Tech.(CSE)-DS-2023/B.Tech.(CSE)-2023

Course Name – Data Handling and Visualization

Course Code - PEC-CSD601B/PEC-CSG601B

(Semester VI)

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) What is the first step in the data analysis process?
 - a) Data Visualization
 - b) Data Collection
 - c) Data Cleaning
 - d) Model Deployment
- (ii) Select the primary purpose of data analytics?
 - a) To store data
 - b) To extract meaningful insights from data
 - c) To delete unnecessary data
 - d) To manually process information
- (iii) What is the main goal of data preprocessing?
 - a) a) To clean and prepare raw data for analysis
 - b) b) To permanently delete old data
 - c) c) To generate reports automatically
 - d) d) To create charts and graphs
- (iv) Why is data-driven decision-making preferred over intuition-based decisions?
 - a) Increases accuracy and objectivity
 - b) Reduces computing costs
 - c) Avoids the need for research
 - d) Makes decision-making slower
- (v) What is a measure of central tendency in data analysis?
 - a) A technique to visualize data
 - b) A statistical measure that represents the center of a dataset
 - c) A method for encrypting data
 - d) A process for storing data
- (vi) Which of the following is a key component of EDA?
 - a) Data encryption
 - b) Data visualization
 - c) Data storage
 - d) Data transmission
- (vii) How does selecting specific rows in a dataset help in analysis?
 - a) It allows for targeted exploration of relevant data points
 - b) It eliminates the need for data storage
 - c) It increases data processing time
 - d) It prevents visualization of trends
- (viii) What does SAS primarily help with in data analytics?

- a) Designing web pages
- b) Advanced statistical modeling and data management
- c) Video rendering
- d) Managing hardware resources
- (ix) How does R differ from Python in data analysis?
 - a) R is primarily used for statistical computing, while Python is more general-purpose
 - b) Python has better built-in visualization tools than R
 - c) R is a proprietary language, while Python is open-source
 - d) R does not support machine learning algorithms
- (x) What is the benefit of integrating Python with Power BI?
 - a) It allows advanced analytics and custom visualizations
 - b) It prevents data manipulation
 - c) It slows down the analysis process
 - d) It restricts data sharing
- (xi) What is one common method to handle missing data?
 - a) Deleting the entire dataset
 - b) Using mean or median imputation
 - c) Replacing missing values with zeros
 - d) Ignoring the missing values
- (xii) What is the best approach when dealing with categorical variables with missing data?
 - a) Replacing missing values with the most frequent category
 - b) Removing all categorical variables
 - c) Converting categorical variables into numerical values
 - d) Ignoring missing values
- (xiii) What does a bar chart typically represent?
 - a) The relationship between two numerical variables
 - b) The distribution of a single categorical variable
 - c) The median value of a dataset
 - d) The correlation between datasets
- (xiv) Why is variance an important statistical measure?
 - a) It measures how much data points deviate from the mean
 - b) It represents the highest value in a dataset
 - c) It determines the median of a dataset
 - d) It counts the number of observations
- (xv) Why would an analyst use a logarithmic scale in data visualization?
 - a) To better visualize data with large variations
 - b) To eliminate outliers
 - c) To ensure all data points have the same value
 - d) To remove unnecessary variables from the dataset

Group-B

(Short Answer Type Questions)

3 x 5=15

2. Create a small Python script to visualize data using a scatter plot. (3)
3. How does Power BI improve business intelligence? (3)
4. What is missing data, and why does it occur in datasets? (3)
5. How does creating secondary variables enhance data analysis? (3)
6. How can EDA help in identifying data quality issues? (3)

OR

Why is sample size important when analyzing a dataset? (3)

Group-C

(Long Answer Type Questions)

5 x 6=30

7. What is the role of R in statistical computing and data analysis? (5)
8. How can Python be used to conduct Exploratory Data Analysis (EDA)? (5)
9. How can dynamic dashboards in Power BI improve business decision-making? (5)
10. Design a data preprocessing workflow for handling missing data, outliers, and feature transformation. (5)
11. What are the common challenges in data cleaning, and how can they be addressed? (5)

12. Compare and contrast descriptive, predictive, and prescriptive analytics with real-world applications. (5)

OR

- How can time-series analysis be used for trend prediction in business and economics? (5)

Library

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