



38019-44241



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

BRAINWARE UNIVERSITY

Term End Examination 2025-2026

Programme – BBA(BA)-Hons-2024/BBA(BA)-Hons-2025

Course Name – Foundations of Data Science

Course Code - BBB20002

(Semester II)

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) Select the primary responsibility of a data scientist.
 - a) Designing web applications
 - b) Extracting insights from data using machine learning and statistics
 - c) Managing physical data centers
 - d) Only collecting raw data
- (ii) Examine which factor is crucial when selecting a sample for statistical inference.
 - a) Sample size and representativeness
 - b) Ignoring data variability
 - c) Collecting data from only one source
 - d) Avoiding randomization
- (iii) Select the programming language widely used in data science for statistical analysis.
 - a) C++
 - b) Java
 - c) Python
 - d) HTML
- (iv) Examine why sampling is essential in data analysis.
 - a) It reduces cost and time while allowing generalization to a population
 - b) It avoids data collection
 - c) It guarantees 100% accuracy
 - d) It eliminates the need for statistical models
- (v) Select the most commonly used language for big data processing in data science.
 - a) SQL
 - b) Swift
 - c) COBOL
 - d) Pascal
- (vi) Cite the condition necessary for a Poisson distribution to be valid.
 - a) The events must occur independently over a fixed interval
 - b) The number of trials must be fixed
 - c) Each trial must have only two possible outcomes
 - d) The probability of success must remain constant
- (vii) Identify the mean of a Poisson distribution if the average event rate is λ .
 - a) λ^2
 - b) λ
 - c) $1/\lambda$
 - d) $\lambda + 1$
- (viii) Cite the type of data that follows a binomial distribution.

- a) Continuous data
- b) Categorical data
- c) Count data with two possible outcomes
- d) Time-series data
- (ix) Identify the key difference between a binomial and Poisson distribution.
 - a) Binomial distribution deals with a fixed number of trials, while Poisson distribution models rare events over time or space
 - b) Both distributions require a fixed probability of success
 - c) Poisson distribution requires two outcomes per trial
 - d) Binomial distribution models the number of events per unit of time
- (x) Identify the primary application of the Chi-Square distribution.
 - a) Testing the independence of categorical variables
 - b) Estimating population means
 - c) Predicting future values
 - d) Analyzing time-series data
- (xi) Choose the best visualization method to evaluate the performance of a classification model.
 - a) Scatter plot
 - b) ROC Curve
 - c) Box plot
 - d) Histogram
- (xii) Choose the most appropriate metric for evaluating a regression model's predictive accuracy.
 - a) Mean Squared Error (MSE)
 - b) F1 Score
 - c) Precision
 - d) Recall
- (xiii) Examine how decision trees help in decision-making.
 - a) They visualize the decision process step by step
 - b) They measure linear relationships
 - c) They eliminate all prediction errors
 - d) They are used only for unsupervised learning
- (xiv) Examine the purpose of a residual plot in model evaluation.
 - a) To visualize the spread of data
 - b) To check for patterns in prediction errors
 - c) To compare different models
 - d) To determine model accuracy
- (xv) Choose the best visualization technique for evaluating a classification model's performance.
 - a) Line plot
 - b) Confusion matrix
 - c) Scatter plot
 - d) Histogram

Group-B

(Short Answer Type Questions)

3 x 5=15

- 2. Define hypothesis testing in statistical inference. (3)
- 3. Define data collection in data science. (3)
- 4. Discuss mutually exclusive and non-mutually exclusive events. (3)
- 5. Explain the role of residual plots in model evaluation. (3)
- 6. Justify the importance of regression analysis in business decision-making. (3)

OR

- Justify the application of machine learning in demand forecasting for retail. (3)

Group-C

(Long Answer Type Questions)

5 x 6=30

- 7. Discuss the differences between Python and Julia for data science. (5)
- 8. Summarize how machine learning enhances recommendation systems (5)
- 9. Define Big Data and Explain Its Characteristics Using the 5Vs Model. (5)
- 10. Define Datafication and Describe Its Role in Business and Society (5)
- 11. Examine the Assumptions of Simple Linear Regression and Their Importance. (5)
- 12. Compare the t-distribution and other probability distributions. (5)

OR

Justify the key properties of the t-distribution and how it compares to the normal distribution.

(5)

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