

BACHELOR OF TECHNOLOGY (CBCS - 2023)
B. Tech. Sem-III Computer Science & Business Systems : WINTER: 2025
SUBJECT: COMPUTATIONAL STATISTICS

Day : Wednesday
Date : 17/12/2025

W-29216-2025

Time : 10:00 AM-01:00 PM
Max. Marks : 60

NB :

1. Assume suitable data, if necessary.
2. Draw neat labelled diagrams WHEREVER necessary.
3. Figures to the right indicate FULL marks for the question.
4. All questions are COMPULSORY.

Q. 1 What is a class in Python? Explain the concept of a constructor and demonstrate different types of constructors (default, parameterized) with an example. (10)

OR

Q. 1 Explain the process of data wrangling. Discuss the significance of data wrangling in data analysis with an example of how it improves the quality of data. (10)

Q. 2 Explain the concept of time ranges, frequencies, and shifting in time series data. How do these operations help in the analysis and manipulation of time-indexed data? (10)

OR

Q. 2 Explain the GroupBy mechanism in data analysis. How does it work, and what are its key advantages for performing split-apply-combine operations on datasets? Provide an example to illustrate your answer. (10)

Q. 3 Explain in detail various assumptions of Multiple Linear Regression model such as Linearity, Collinearity, Outliers, Autocorrelation, Homoscedasticity and Normality. (10)

OR

Q. 3 Derive the formula for Multivariate Normal Distribution where variables are independent on each other. (10)

Q. 4 Discuss the key assumptions of multivariate regression analysis. Explain how do violations of these assumptions affect the validity of the model. (10)

OR

Q. 4 How to classify a new observation using Fisher Linear Discriminant Analysis. Explain in detail. (10)

Q. 5 What is factor analysis? Explain its purpose and how it differs from Principal Component (10)

Analysis (PCA). Discuss the key assumptions underlying factor analysis.

OR

Q. 5 How to determine the optimal number of principal components to retain in PCA? Explain with reference to variance explained and methods like the scree plot. (10)

Q. 6 Explain the different kinds of distance metrics used in clustering to measure proximity between data points. Give examples of when each distance metric is most appropriate. (10)

OR

Q. 6 What is a dendrogram? Explain how it is used in hierarchical clustering with an example to demonstrate how clusters are formed and visualized. (10)

* * * * *

171225-m-coe-mumbai