

**BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)**

**CLASS: MBA
BRANCH: MBA**

**SEMESTER : IV
SESSION : SP/2025**

SUBJECT: MT550 MULTIVARIATE DATA ANALYSIS

TIME: 3 Hours

FULL MARKS: 50

INSTRUCTIONS:

1. The question paper contains 5 questions each of 10 marks and total 50 marks.
 2. Attempt all questions.
 3. The missing data, if any, may be assumed suitably.
 4. Before attempting the question paper, be sure that you have got the correct question paper.
 5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.
-

- | | | CO | BL | | | | | | | | | | | | | | | |
|---|-----|----|----|--------|----|---|----|----|---------|----|----|----|---|-------|---|----|---|---|
| Q.1(a) Define the following terms: 1) Sample mean vector 2) dummy variables | [5] | 1 | 1 | | | | | | | | | | | | | | | |
| Q.1(b) State the Gauss-Markov theorem and explain its implications for the properties of the Ordinary Least Squares (OLS) estimator. What assumptions must hold for the theorem to be valid? | [5] | 1 | 1 | | | | | | | | | | | | | | | |
| Q.2(a) What is Confirmatory Factor Analysis (CFA), and how does it differ from EFA? | [5] | 2 | 2 | | | | | | | | | | | | | | | |
| Q.2(b) Discuss the different types of clustering algorithms. How can Cluster Analysis be used to identify customer segments or product categories? | [5] | 2 | 1 | | | | | | | | | | | | | | | |
| Q.3(a) To test the significance of variation in the retail prices of a commodity in three principal cities, Mumbai, Kolkata, and Delhi, four shops were chosen at random in each city and the prices who lack confidence in their mathematical ability observed in rupees were as follows: | [5] | 3 | 4 | | | | | | | | | | | | | | | |
| <table border="1" style="margin: auto; border-collapse: collapse;"><tbody><tr><td style="padding: 2px;">Kanpur</td><td style="padding: 2px;">15</td><td style="padding: 2px;">7</td><td style="padding: 2px;">11</td><td style="padding: 2px;">13</td></tr><tr><td style="padding: 2px;">Lucknow</td><td style="padding: 2px;">14</td><td style="padding: 2px;">10</td><td style="padding: 2px;">10</td><td style="padding: 2px;">6</td></tr><tr><td style="padding: 2px;">Delhi</td><td style="padding: 2px;">4</td><td style="padding: 2px;">10</td><td style="padding: 2px;">8</td><td style="padding: 2px;">8</td></tr></tbody></table> | | | | Kanpur | 15 | 7 | 11 | 13 | Lucknow | 14 | 10 | 10 | 6 | Delhi | 4 | 10 | 8 | 8 |
| Kanpur | 15 | 7 | 11 | 13 | | | | | | | | | | | | | | |
| Lucknow | 14 | 10 | 10 | 6 | | | | | | | | | | | | | | |
| Delhi | 4 | 10 | 8 | 8 | | | | | | | | | | | | | | |
| Q.3(b) Do the data indicate that the price in the three cities are significantly different? The table value of F for $df_1 = 2$, $df_2 = 9$, and $\alpha = 5\%$ level of significance is 4.26. What are the key advantages of SEM over traditional regression analysis in terms of addressing complex relationships and modeling latent variables? | [5] | 3 | 2 | | | | | | | | | | | | | | | |
| Q.4(a) Explain the concept of canonical variates. How do they relate to the separation of groups in discriminant analysis? | [5] | 4 | 3 | | | | | | | | | | | | | | | |
| Q.4(b) Explain what a binary logit model is and how it differs from linear regression. What are the main advantages of using a logit model for binary outcomes? | [5] | 4 | 2 | | | | | | | | | | | | | | | |
| Q.5(a) Explain the concept of Functional Principal Component Analysis (FPCA) and its role in data reduction and feature extraction for functional data. How could FPCA be used to identify key drivers of sales performance in a retail business? | [5] | 5 | 4 | | | | | | | | | | | | | | | |
| Q.5(b) How can functional data analysis be used in the healthcare industry to analyze patient data? What are the potential benefits and challenges? | [5] | 5 | 4 | | | | | | | | | | | | | | | |

:::01/05/2025:::E