



25047656

QP CODE: 25047656

Reg No :

Name :

M.Sc DEGREE (CSS) EXAMINATION, NOVEMBER 2025

Third Semester

MSc Statistics with Data Science

**Core Course - ST040301 - MULTIVARIATE ANALYSIS AND STATISTICAL
TECHNIQUES FOR DATA MINING**

2020 ADMISSION ONWARDS

03F36E87

Time: 3 Hours

Weightage: 30

Part A (Short Answer Questions)

*Answer any **eight** questions.*

Weight 1 each.

1. What is data mining? Briefly explain the preprocessing steps of data mining?
2. Explain data discretization.
3. Define communality and specific variance.
4. State the importance of rotation in Factor analysis.
5. When will canonical correlation reduces to bivariate correlation?
6. State the condition which leads to quadratic discriminant function in the case of classification of two normal populations .
7. Give the Fishers discriminant analysis classifier.
8. Distinguish between hierarchical clustering and non-hierarchical clustering.
9. Give the two way MANOVA Model.
10. Explain sphericity test.

(8×1=8 weightage)

Part B (Short Essay/Problems)

*Answer any **six** questions.*

Weight 2 each.

11. Explain evolution and deviation analysis.





12. Obtain principal component from standardized variables.
13. Let $\Sigma = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 4 \end{bmatrix}$. Determine the Principal components. What is the proportion of variance explained by the first principal component. Also find the specific variances.
14. State holdout procedure for estimating the efficiency of discriminant analysis.
15. Define single, complete and average linkage methods.
16. Give the test procedure for testing whether mean values are same when population covariance matrix is given.
17. Briefly discuss one way MANOVA.
18. Define Pillai's trace statistic and Roy's maximum root Statistics.

(6×2=12 weightage)

Part C (Essay Type Questions)

Answer any **two** questions.

Weight **5** each.

19. Explain the tests considered in Profile analysis.
20. Obtain the minimum expected cost of misclassification rule with equal misclassification cost in the case with more than two populations.
21. Present the motivation, definition and derivation of Fisher's (multiple) discriminant functions.
22. Describe how will you test the equality of co-variance matrices of q multivariate normal populations.

(2×5=10 weightage)

