



Mahatma Gandhi University

Kottayam

Programme	BSc (Hons) Statistics					
Course Name	Introduction to Multivariate Analysis					
Type of Course	DSE					
Course Code	MG5DSESTA300					
Course Level	300 - 399					
Course Summary	To provide the students with knowledge of the statistical concepts of multivariate data analysis and their basic methodology.					
Semester	5	Credits			4	Total Hours
Course Details	Learning Approach	Lecture	Tutorial	Practical	Others	
		3		1		75
Pre-requisites						

COURSE OUTCOMES (CO)

CO No.	Expected Course Outcome	Learning Domains	PO No
1	Apply multivariate testing procedures to real life datasets.	A	2
2	Apply MDS and PCA for dimension reduction.	A	2
3	Apply Factor analysis for identification of latent variables.	A	2
4	Classify the multivariate observations into groups using classification methods.	E	2
5	Identify patterns in data using cluster and correspondence analyses.	E	2

**Remember (K), Understand (U), Apply (A), Analyse (An), Evaluate (E), Create (C), Skill (S), Interest (I) and Appreciation (Ap)*

COURSE CONTENT

Content for Classroom Transaction (Units)

	Course Description	Hours	CO No.
Module 1	Dimension Reduction Techniques	15	
1.1	Hotelling's and Mahalanobis statistics, their properties, inter-relationships and uses. Canonical variates and canonical correlation, use, estimation, and computation.	6	1
1.2	Profile Analysis and the associated tests.	3	1
1.3	Multidimensional Scaling, Principal Component Analysis (PCA)- Method of extraction-properties, the associated tests.	6	2
Module 2	Latent Variable Identification	15	
2.1	Factor Analysis-Types- Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA-Orthogonal Model.	9	3
2.2	Estimation of factor loadings, factor rotations.	6	3
Module 3	Statistical Machine Learning	15	
3.1	Bayes' Classifier, Fisher's linear discriminant function	3	4
3.2	Support Vector Machine, PCA approach.	3	4
3.3	Classification trees and K-Nearest Neighbors (KNN) algorithm.	3	4
3.4	Cluster Analysis: proximity measures	2	5
3.5	Hierarchical and non-hierarchical methods.	2	5
3.6	Correspondence Analysis.	2	5
4	Practical using R/Python	30	
4.1	A record with minimum 10 Problems has to be submitted	30	1,2,3,4,5.
Module 5	Teacher Specific Content.		

Teaching and Learning Approach	Classroom Procedure (Mode of transaction) Direct Instruction: Brainstorming lecture, E-learning, Interactive Instruction, Seminar, Group Assignments, Authentic learning, Presentation by students by group.
---------------------------------------	--

Assessment Types	<p>MODE OF ASSESSMENT</p> <p>A. Continuous Comprehensive Assessment (CCA)</p> <p>Practical Formative assessment : 15 marks Lab involvement, Practical Record, Viva voce</p> <p>Theory Formative assessment : 10 marks Quiz, Assignments, Seminar. Summative assessment : 15 marks Written tests</p> <p style="text-align: center;">Theory: 25 marks (Total)</p>
	<p>B. End Semester Evaluation (ESE) Theory: 50 marks (Duration : 1.5 Hrs)</p> <p>i) Short answer type questions: Answer any 7 questions out of 10 (7*2=14).</p> <p>ii) Short essay type questions: Answer any 4 questions out of 6 (4*6=24).</p> <p>iii) Essay type questions: Answer any 1 question out of 2 (1*12=12).</p> <p>Practical: 35 marks Problem solving skills: 30 marks Record: 5 marks</p>

References:

1. Johnson, R.A. and Wichern, D.W. (2013). Applied Multivariate Statistical Analysis, 6th Edition. Pearson Education.
2. Vander Plas, J. (2022). Python Data Science Handbook: Essential Tools for Working with Data. 2nd Edition. Shroff Publishers & Distributors Pvt. Ltd
3. Brian Everitt, Torsten Hothorn (2011). An Introduction to Applied Multivariate Analysis with R, Springer New York, NY.

Suggested Readings:

1. Anderson, T.W.(2009).An Introduction to Multivariate Statistical Analysis, 3rd Edition, John Wiley.
2. VanderPlas, J. (2022). Python Data Science Handbook: Essential Tools for Working with Data. 2nd Edition. Shroff Publishers & Distributors Pvt. Ltd.
3. Rencher, A.C.(1998). Multivariate Statistical Inference and Applications,1st Edition, Wiley-Interscience.
4. Seber G. F. (2004). Multivariate Observations. 1st Edition, John Wiley & Sons.