Course title: Multivariate Data Analysis				
Course code: NRE 112	**No. of credits:** 3	**L-T-P:** 28-14-0		**Learning hours:** 42
Pre-requisite course code and title (if any): NRE 115 Environmental Statistics				
Department: Energy and Environment				
Course coordinator(s):		**Course instructor(s):** Kaushik Roy Chowdhury		
Contact details:				
Course type: Elective		**Course offered in:** Semester 2		

Course Description

Large amount of data is collected on many different variables across disciplines in order to understand the underlying process(es). The multivariate analysis of data deals with examining interrelationship between three or more equally important variables or explaining of variation in, usually one (or more than one) dependent variable(s) on the basis of two or more independent (explaining) variables. With the availability of inexpensive, fast and efficient computing resources and statistical packages there has been a growth in the application of these techniques. This course introduces the student to various multivariate data analysis tools. The focus is on cross-disciplinary application of these techniques.

Course objectives

- Introduce the language of multivariate data analysis
- Understand the characteristics of multivariate quantitative research, including strengths and weaknesses
- Understand the principles and characteristics of the multivariate data analysis techniques

Course content

SNo	Topic	L	T	P
1.	**Introduction** Basic multivariate statistics–mean, variance, ovariance, correlation, linear combination of variables, geometric concepts, distances	4	2	
2.	**Interdependence methods** Principal component analysis, factor analysis, cluster analysis, correspondence analysis, multidimensional scaling	12	6	
3.	**Dependence methods** Multiple regression models, logistic regression canonical correlation, discriminant analysis.	12	6	
	Total	**28**	**14**	

Evaluation criteria

- 2 minor test: 15% each
- Assignment + viva voce: 20
- Major test: 50

Learning outcomes

- distinguish between dependence and interdependence methods in multivariate data analysis
- identify the most appropriate statistical techniques for a multivariate dataset
- carry out and apply commonly used multivariate data analysis techniques, and interpret results
- use statistical software packages for the analysis of multivariate data

Pedagogical approach

Materials

Required text

Afifi A., May S. and Clark V.A. (2012) *Practical Multivariate Analysis*, CRC Press, Taylor &Francis, Boca Raton.

Johnson R.A. and Wichern D.W. (2002) *Applied Multivariate Statistical Analysis*, Prentice Hall of India Pvt Ltd., New Delhi.

Sharma S. (1996) *Applied Multivariate Techniques*, John Wiley and Sons, New York.

Suggested readings

Alt M. (1990) *Exploring Hyperspace–A Non-mathematical Explanation of Multivariate Analysis,* McGraw-Hill Book Company, New York

Everitt B.S. and Dunn G. (2001), *Applied Multivariate Data Analysis*, Arnold, London.

Haan C.T. (1977) *Statistical Methods in Hydrology*, The Iowa State University Press/Ames.

Harris R.J. (1985) *A Primer in Multivariate Statistics*, Academic Press, New York.

Manly B.F.J. (1994) *Multivariate Statistical Methods–A Primer*, Chapman and Hall, London.

Case studies
Websites

Journals

Applied Statistics
Biometrics
Biometrika
Environmental and Ecological Statistics
Environmetrics
Journal of the American Statistical Association
Psychometrika
Statistical Science
Technometrics
The American Statistician

Additional information (if any)

Student responsibilities

The students are expected to submit assignments in time and come prepared with readings when provided.