

## Liverpool John Moores University

Title: MULTIVARIATE ANALYSIS AND DATA MINING  
Status: Definitive  
Code: **6006STATS** (117484)  
Version Start Date: 01-08-2018  
Owning School/Faculty: Applied Mathematics  
Teaching School/Faculty: Applied Mathematics

Team	Leader
Sandra Ortega Martorell	Y
Ivo Siekmann	

**Academic Level:** FHEQ6      **Credit Value:** 24      **Total Delivered Hours:** 75  
**Total Learning Hours:** 240      **Private Study:** 165

### Delivery Options

Course typically offered: Standard Year Long

Component	Contact Hours
Lecture	22
Practical	20
Tutorial	30

**Grading Basis:** 40 %

### Assessment Details

Category	Short Description	Description	Weighting (%)	Exam Duration
Technology	AS1	Minitab based coursework	20	
Report	AS2	Neural network based coursework	20	
Exam	AS3	Examination	60	3

### Aims

*To enable the student to explore the structure of multidimensional data sets.  
To introduce the student to inferential procedures using multivariate data.  
To provide the student with familiarity with linear and flexible methods for regression*

and classification.

## Learning Outcomes

After completing the module the student should be able to:

- 1 Carry out an exploratory numerical and graphical analysis of a set of multivariate data.
- 2 Recognize situations in which a multivariate approach is required and carry out the appropriate inferential procedures.
- 3 Classify future multivariate observations into one of a number of known populations.
- 4 Apply linear methods.
- 5 Apply neural network algorithms.
- 6 Report their conclusions in an appropriate manner.

## Learning Outcomes of Assessments

The assessment item list is assessed via the learning outcomes listed:

Minitab	1	2			
Neural networks	5				
Examination	1	2	3	4	6

## Outline Syllabus

*Graphical display and numerical summary of multivariate data.*

*Investigation of the dependence among variables.*

*Discrimination and prediction. Error rate estimation.*

*Hypothesis construction and testing. Use of simultaneous confidence intervals.*

*Principal Components Analysis.*

*Use of R for data exploration, parameter estimation and significance testing.*

*Stages of the data analysis using data mining, including how to create and evaluate models*

*Neural networks, including backpropagation, gradient descent, the momentum method*

*Clustering with k-means and k-medoids methods*

*Blind Source Separation, including Independent Component Analysis and Non-negative Matrix Factorisation*

*Self-Organising Maps*

*Introduction to Big Data*

## Learning Activities

Lectures, tutorials, laboratory sessions, directed reading, coursework preparation and revision for examinations.

## **Notes**

Multivariate data consist of observations taken on several variables from each experimental unit. The special problems associated with data of this type will be covered in this module. The course extends beyond linear methods into adaptive non-linear modeling.