

Making Sense of Multivariate Data

This course covers the excellent features in XLSTAT for investigating, visualising and modelling data sets with measurements on many variables. Depending on the application your focus may be on the objects measured (customers, products, physical samples etc.) or the inter relationships between the variables themselves. There are multivariate techniques available to suit both these aims.

Mathematical details are kept to a minimum and we focus on the interpretation of the outputs from XLSTAT and illustrate applications with case studies using data from consumer tests, instrumental analysis, manufacturing processes and biological data.

We aim to explain the objectives of the techniques and to send you away with a better understanding of which technique to use when.

There is plenty of opportunity for practice using XLSTAT through “hands on exercises” for which annotated solutions are provided and delegates are welcome to bring their own datasets to try the techniques.

Day 1

- Multivariate data – correlation and association measures
 - Different data types, how to measure their interrelationships
 - Covariance, correlation and Chi-squared association
- Principal Component Analysis
 - How it works
 - Finding your way through the maze of statistics
 - Applications to data reduction and data display
 - How to interpret the plots
 - Adding supplementary data (a great XLSTAT option)
 - Comparison with Factor Analysis
- Cluster Analysis
 - Objectives and when to use the technique
 - Measuring proximity –similarity and dissimilarity
 - Hierarchical Clustering
 - Disjoint Clustering using Kmeans
 - Comparison of the approaches
 - Evaluating the solution

Day 2

- Correspondence Analysis
 - How it works
 - When to use it
 - Interpretation of the outputs
 - Reading the map
 - Extensions to Joint and Multiple Correspondence Analysis
 - Typical applications
- Factor Analysis
 - How it works and what it does
 - Comparison with Principal Components Analysis

- Interpretation of output
- Typical applications
- Multiple Regression
 - Review of classical regression techniques
 - Multiple regression
 - Automatic variable selection methods
 - Diagnostic measures –evaluating the fit of the model
 - Problems and possible pitfalls
- Discriminant Analysis
 - Aims of technique
 - Discriminant Functions
 - Evaluating success – interpreting the confusion matrix
 - Visualisation using canonical variates
- Course Review
 - Review of opportunities with each technique
 - Which method to use - when

The course will be held at a state of the art training venue in the City of London.

Computers will be provided so there is no need to bring your laptop (unless you want to work on some of your own data)