



Multivariate Data Analysis

WHY MULTIVARIATE DATA ANALYSIS?

The massive amounts of (non-designed) collected data, often stored without further analysis, might contain valuable information about wanted and unwanted variation in process factors and product properties. A multivariate approach might reveal the cause of the unwanted variation or phenomena, without any additional experiments or measurements. Taking into account all available information will lead to insights in the often complex interplay of many factors.

COURSE SET-UP

During day 1 qualitative aspects of multivariate data analysis will be treated: exploring the data, searching for correlations, clusters, outliers, ...

In day 2 we come to the model building part: searching for relations between groups of variables. Emphasis will be put on correctly applying and interpreting the different techniques, and not on underlying theory. The course matter can immediately be applied with real-life exercises on PC.

COURSE OBJECTIVE

Multivariate analysis comprises a broad gamma of techniques, but at the same time contains an equally broad gamma of pitfalls. Breaking down the barriers towards multivariate analysis and smoothing the path towards expertise building while at the same time making the participants aware of the problems that arise, are considered to be the main objectives of this course.

At the end of the course participants will be able to select the proper technique to solve a number of problems, analyse the data and correctly interpret the results.

INTENDED AUDIENCE AND PRIOR KNOWLEDGE

This course will be of great help to anyone who is frequently faced with large data tables and who is not familiar with multivariate techniques, who can not make the appropriate choice, or who does not feel confident when trying to interpret the results of these methods.

Prior knowledge is not required.

PRACTICAL

Each course day will be held from 9 am to about 5 pm. The course dates and fees are listed on the attached application form. The course fee includes handouts and lunches.

To apply, send us back the application form, or apply on-line, at least 20 days before the start of the course .



COURSE CONTENTS

Day 1: Exploratory multivariate analysis

- Visualisation of big datasets
- Principal Component Analysis (PCA)
- Cluster analysis: searching for groups of similar samples

Day 2: Quantitative analysis: in search of cause-effect relations

- Multiple Linear Regression (MLR) with uncorrelated variables
- Multiple Linear Regression (MLR) with correlated variables
 - Stepwise regression
 - The collinearity problem
 - An overview of the pitfalls
- Principal Component Regression (PCR)
- Partial Least Squares (PLS)
 - Interpretation of PCR and PLS models
 - Validation of regression models
 - Detection of outliers and non-linearities
 - Prediction with regression models
- Some alternatives

Day 3: Quantitative analysis: the sequel + specific applications

- Feasibility study: does a quantitative analysis make sense?
- Classification (supervised pattern recognition): predicting class membership
 - Classification rules
 - Linear Discriminant Analysis (LDA)
 - Soft Independent Modeling of Class Analogy (SIMCA)
 - PLS-DA
- Specific applications:
 - QSAR / QSPR (Quantitative Structure Activity / Property Relations)
 - Multivariate SPC (M-SPC)
 - Principal Properties Design
 -