



# Multivariate Datenanalyse

## WARUM MULTIVARIATE DATENANALYSE?

Die großen Mengen (ohne Versuchsplanung) erfasster Daten, die häufig ohne weitere Analyse gespeichert werden, könnten wertvolle Informationen über erwünschte und unerwünschte Variationen in Prozessfaktoren und Produkteigenschaften enthalten. Eine multivariate Methode könnte ohne zusätzliche Experimente oder Messungen die Ursache der unerwünschten Variation oder Phänomene aufzeigen. Die Berücksichtigung der gesamten verfügbaren Informationen verleiht Einblicke in das oftmals komplexe Zusammenspiel vieler Faktoren.

## KURSVERLAUF

Am Tag 1 werden qualitative Aspekte der multivariaten Datenanalyse behandelt: Untersuchung der Daten, Auffinden von Korrelationen, Clustern, Ausreißern, ...

Am Tag 2 folgt der Modellerstellungsteil: Suche nach Zusammenhängen zwischen Gruppen von Variablen. Der Schwerpunkt liegt auf der korrekten Anwendung und Interpretation der verschiedenen Techniken und nicht auf der zugrunde liegenden Theorie. Der Kursinhalt kann sofort bei realen Übungen am PC angewendet werden.

## ZIELSETZUNG DES KURSES

Die multivariate Analyse umfasst ein breites Spektrum von Techniken, beinhaltet aber gleichzeitig ein gleich großes Spektrum an Fallstricken. Als primäre Zielsetzungen sollen in diesem Kurs Schranken gegenüber der multivariaten Analyse abgebaut und der Weg zur Erwerbung einer Fachkompetenz geebnet werden, während den TeilnehmerInnen gleichzeitig die damit verbundenen Probleme bewusst gemacht werden.

Am Ende des Kurses können die TeilnehmerInnen die richtige Technik zur Lösung einer Reihe von Problemen wählen, die Daten analysieren und die Ergebnisse interpretieren.

## ZIELGRUPPE UND VORKENNTNISSE

Dieser Kurs ist für Personen hilfreich, die häufig mit großen Datentabellen konfrontiert und nicht mit multivariaten Techniken vertraut sind, nicht die richtige Wahl treffen können oder sich die Interpretation der Ergebnisse dieser Methoden nicht zutrauen.

Keine Vorkenntnisse erforderlich.



## KURSinHALT

---

### Tag 1: Explorative multivariate Analyse

- Visualisierung von großen Datensätzen
- Principal Component Analysis (PCA)
- Cluster-Analyse: Ermittlung von Gruppen gleicher Proben

### Tag 2: Quantitative Analyse: die Suche nach Kausalbeziehungen

- Multiple Lineare Regression (MLR) mit nicht korrelierten Variablen
- Multiple Lineare Regression (MLR) mit korrelierten Variablen
  - Stufenweise Regression
  - Das Kollinearitätsproblem
  - Überblick über die Fallstricke
- Principal Component Regression (PCR)
- Partial Least Squares (PLS)
  - Interpretation von PCR und PLS Modellen
  - Validierung von Regressionsmodellen
  - Detektion von Ausreißern und Nicht-Linearitäten
  - Vorhersage mit Regressionsmodellen
- Einige Alternativen

### Tag 3: Quantitative Analyse: die Folge + spezifischen Anwendungen

- Machbarkeitsstudie: ergibt eine quantitative Analyse Sinn?
- Klassifizierung (überwachte Mustererkennung): Vorhersage der Klassenzugehörigkeit
  - Klassifizierungsregeln
  - Linear Discriminant Analysis (LDA)
  - Soft Independent Modeling of Class Analogy (SIMCA)
  - PLS-DA
- Spezifische Anwendungen:
  - QSAR / QSPR (Quantitative Structure Activity / Property Relations)
  - Multivariate SPC (M-SPC)
  - Principal Properties Design
  - .....

## ALLGEMEINE INFORMATIONEN

---

Jeder Kurstag findet von 9:00 h bis etwa 17:00 h statt. Die Kursgebühr und die Kursdaten finden Sie in dem beiliegenden Anmeldeformular. In den Kursgebühren sind Unterlagen und Mittagessen enthalten. Für die Anmeldung retournieren Sie bitte das Anmeldeformular spätestens 30 Tage vor Kursbeginn.