



GESELLSCHAFT DEUTSCHER CHEMIKER

## Präsenzkurs: Einsatz der statistischen Software R: Grundlagen, Data-Mining und maschinelles Lernen (643/21)

Mit Hygienekonzept

Optimierte Datenauswertungen in der  
chemischen Forschung und Produktion

Prof. Dr. Bernard Ludwig

- Einführung in die mächtige und freie statistische Software R: Aufbau, Objekte & Funktionen
- Einsatz von R für
  - Signifikanztests, Regressionen & Varianzanalysen
  - Chemometrie
  - Data Mining & maschinelles Lernen



20. – 21. September 2021 · Frankfurt am Main

### ZIEL

In vielen chemischen Bereichen in Forschung und Produktion werden Daten erhoben – u.a. für Qualitätskontrollen, Produktionsoptimierungen, Verkaufsprognosen und Mustererkennungen für Automatisierungen. Im Rahmen des Industrie 4.0-Konzepts ist zudem ein Verständnis künstlicher Intelligenz und damit von Data-Mining-Verfahren und maschinellem Lernen bedeutsam – u.a. für Optimierungen vernetzter Produktionssysteme. Ziel der Veranstaltung ist es, dass die Teilnehmer Auswertungen mit der mächtigen und freien statistischen Software R erlernen, die ein breites Methodenspektrum für Auswertungen bietet.

### INHALT

Im Teil I wird eine Einführung in R gegeben. Die Teilnehmer werden in die Lage versetzt, Zielvariablenoptimierungen (z.B. Produktausbeute) in Abhängigkeit kategorialer (z.B. Produktionsverfahren) und quantitativer Variablen (z.B. Temperatur) und Signifikanztests (z.B. Reinheitsvergleich) durchzuführen und chemometrische Verfahren anzuwenden. R wird eingesetzt für:

- Deskriptive und inferentielle Statistik (u.a. klassische Tests, Varianzanalyse)
- Explorative Statistik (u.a. statistische Modellierung und Regressionen)
- Chemometrie (u.a. Hauptkomponentenanalyse und Partial Least Squares-Regression)

Im Teil II werden der Einsatz von R in Data-Mining-Verfahren und maschinellem Lernen anhand von Fallbeispielen vermittelt. Behandelt werden Klassifikations- und Regressionsprobleme bei unüberwachtem und überwachtem Lernen. Vorgestellte Methoden sind:

- Cluster- und Faktorenanalysen
- Random Forest und Support Vector Machine-Klassifikationen und Regressionen
- Neuronale Netze

### ZIELGRUPPE

Die Veranstaltung richtet sich an Chemiker, Pharmazeuten, Chemieingenieure und chemisch-technische Mitarbeiter – in Industrie und Hochschule – die an der Datenauswertung beteiligt sind, ausgehend von klassischen Signifikanztests (z.B. Einsatz von R für den Vergleich verschiedener Produktchargen mithilfe von Signifikanztests) über Standard-Regressionen und Varianzanalysen (z.B. Ausbeutenmaximierung) zu neueren Verfahren des Data-Mining und maschinellen Lernens (z.B. für verbesserte Mustererkennungen und Regressionen bei multivariaten Datensätzen).

### VORKENNTNISSE

Grundlagenkenntnisse in einfachen statistischen Auswertungen in der Chemie und in EDV-Anwendungen sind wünschenswert. R-Kenntnisse werden nicht vorausgesetzt.

### TEILNEHMERZAHL

maximal 20 Personen

### MONTAG, 20. SEPTEMBER 2021

- 10.00 Begrüßung der Teilnehmer und Vorstellung
- 10.15 Einführung in R: Datentypen, Vektoren, Datenrahmen, Einlesen von Daten und klassische Tests
- 12.30 Mittagspause
- 13.30 Einsatz von R für Zielvariablenoptimierungen: Regressionen und Varianzanalysen mit Laptop-Übungen
- 15.30 Kaffeepause
- 16.00 Klassifikationsprobleme in der Forschung und Optimierungsprobleme in der Produktion: u. a. Hauptkomponentenanalyse und PLS-Regression mit Laptop-Übungen
- 18.15 Voraussichtliches Ende des ersten Veranstaltungstages

### DIENSTAG, 21. SEPTEMBER 2021

- 9.00 Data Mining und maschinelles Lernen: Wozu braucht man das? Allgemeine Informationen zu Klassifikationen und Regressionen
- 10.00 Kaffeepause
- 10.15 Einsatz verschiedener R-Pakete für Mustererkennungen, Prognosen und Optimierungen
- 12.30 Mittagspause
- 13.30 Data Mining & maschinelles Lernen I: Random Forest mit Laptop-Übungen
- 14.30 Kaffeepause
- 15.00 Data Mining & maschinelles Lernen II: künstliche neuronale Netze und Support Vector Machine mit Laptop-Übungen
- 17.00 Teilnehmer-Feedback und Verabschiedung
- 17.30 Voraussichtliches Ende der Veranstaltung

Änderungen und Ergänzungen vorbehalten

Obwohl im Text häufig nur von Chemikern, Teilnehmern etc. die Rede ist, sind damit selbstverständlich alle Geschlechter gemeint.

## ANMELDUNG

Melden Sie sich bitte online unter [www.gdch.de/64321](http://www.gdch.de/64321) bis zum 23.8.2021 (Anmeldeschluss) bei der Gesellschaft Deutscher Chemiker e.V. (GDCh) an:



Lena Rubner  
Fortbildungsorganisation

T: +49 69 7917-364  
l.rubner@gdch.de  
[www.gdch.de/fortbildung](http://www.gdch.de/fortbildung)

## GEBÜHREN

GDCh-Mitglied € 990,-  
Nichtmitglied € 1.070,-

Die Gebühren sind einschließlich Begleitmaterial und GDCh-Zertifikat, Mittagessen, Kaffeepausen- und Konferenzgetränken, ausschließlich Unterkunft zu verstehen. Sie unterliegen nicht der Mehrwertsteuerpflicht (Steuerbefreiung nach § 4 Nr. 21. a) bb) UStG).

Die AGB finden Sie unter [www.gdch.de/teilnahme](http://www.gdch.de/teilnahme).

## VERANSTALTUNGSORT

Maritim Hotel Frankfurt  
Theodor-Heuss-Allee 3  
60486 Frankfurt am Main  
T: +49 69 7578-0  
info.fra@maritim.de  
[www.maritim.de](http://www.maritim.de)

## ANFAHRT

Die Deutsche Bahn bietet attraktive Konditionen für Ihre Anreise zu GDCh-Veranstaltungen an. Informationen erhalten Sie unter [www.gdch.de/bahn](http://www.gdch.de/bahn).

## UNTERKUNFT

Für die Teilnehmer haben wir im Veranstaltungshotel unter dem Stichwort „GDCh 643/21“ ein begrenztes Zimmerkontingent zu Sonderkonditionen reserviert. Dieses Kontingent gilt bis zum 23.8.2021. Bitte wenden Sie sich direkt an das Hotel (Adresse und Telefonnummer siehe „Veranstaltungsort“).

Weitere Unterkünfte erfragen Sie bitte bei:  
Tourismus+Congress GmbH  
Kaiserstraße 56  
60329 Frankfurt am Main  
T: +49 69 21 230808  
info@infofrankfurt.de  
[www.frankfurt-tourismus.de](http://www.frankfurt-tourismus.de)

Wir weisen ausdrücklich darauf hin, dass die Haftung für bestellte und nicht abgenommene Zimmer beim Besteller liegt.

## HINWEIS AUF WEITERE VERANSTALTUNGEN

595/21 **Präsenzkurs: Rheologische Charakterisierung von Emulsionen und Suspensionen**  
Leitung: Prof. Dr. Karl-Heinz Jacob  
27. – 28. September 2021 · Nürnberg

001/21 **Pigmente – aktueller Stand**  
Leitung: Dr. Carsten Handrosch  
8. – 11. November 2021

511/21 **Hybrid: Die Qualitätssysteme GMP (Gute Herstellungspraxis) und GLP (Gute Laborpraxis) im Überblick – Ein Leitfaden der Guten Praxis**  
Kursmodul zum Geprüften Qualitätsexperten GxP (GDCh)  
Leitung: Dr.-Ing. Barbara Pohl  
22. November 2021 · Frankfurt am Main und Online

517/21 **Online-Kurs: Qualitätsmanagement im analytischen Labor**  
Richtlinienkonformität und Kompetenzerhalt:  
technische Grundlagen qualitätsgerechter Laborarbeit  
(gemeinsam veranstaltet mit EUROLAB/Deutschland)  
Leitung: Dr. Michael Koch  
23. – 24. November 2021 · Online

907/21 **Chemical Development and Scale-Up in the Fine Chemical and Pharmaceutical Industries**  
Leitung: Dr. Will Watson  
November 23 - 25, 2021

991/21 **Online-Kurs: Patent-Know-how für Chemiker**  
Leitung: PA Dr. Hans-Peter Jönsson  
25. November 2021 · Online

588/21 **Online-Kurs: Datenmanagement und regulatorische Anforderungen zur Erstellung und Pflege von Sicherheitsdatenblättern**  
Am Beispiel von Software-Lösungen  
Leitung: Dr. Thorben Bonarius  
29. November 2021 · Online

### INHOUSE-KURSE Lokal oder digital

Individuell, effizient, zeit- und kostensparend –  
nutzen Sie das Expertenwissen und unser Know-how  
als langjähriger Seminaranbieter, auch für Ihre Inhouse-Kurse  
vor Ort oder digital.

Ihre Ansprechpartnerin: Melanie Sakarya  
T: +49 69 7917-331 oder fb@gdch.de

## LEITUNG



Prof. Dr. Bernard Ludwig  
Universität Kassel  
Fachgebiet Umweltchemie

Bernard Ludwig leitet seit 2002 das Fachgebiet Umweltchemie an der Uni Kassel. Seine Forschungen behandeln u.a. umweltchemische, bodenkundliche und spektroskopische Themen, bei denen statistische und Data-Mining-Verfahren und maschinelles Lernen unter Verwendung der Software-Pakete R und SAS eingesetzt werden. Der Fokus vieler seiner in internationalen Fachzeitschriften publizierten Studien liegt auf Methodenvalidierungen und -optimierungen. In der Lehre vermittelt er Methoden der angewandten Statistik, Data-Mining und maschinelles Lernen für Datenauswertungen.

## REFERENT

Prof. Dr. Bernard Ludwig Leiter des Fachgebiets Umweltchemie,  
(siehe Leitung) Universität Kassel

## STOFFVERMITTLUNG

Statistisches Hintergrundwissen wird in verständlicher Form erläutert und Anwendungen der R-Software werden anhand von Fallbeispielen aus Forschung und Produktion gezeigt. Ein Schwerpunkt liegt auf den Laptop-Übungen, sodass die Teilnehmer in die Lage versetzt werden, R für eigene Fragestellungen einzusetzen.

## BEGLEITMATERIAL

Die Teilnehmer erhalten während des Kurses schriftliches Begleitmaterial sowie nach erfolgreicher Teilnahme ein GDCh-Zertifikat.