



WWU WEITER BILDUNG



› KONZEPT

Die statistische Programmiersprache R hat sich zu einem beliebten und vielfältigen Werkzeug entwickelt. Die Open Source Software (www.r-project.org) bietet seinen Anwendern zahlreiche Vorteile: R erfordert keine Lizenzen/ist gratis, läuft auf allen gängigen Betriebssystemen (Windows, Mac OS X, Linux), wird regelmäßig aktualisiert, besitzt eine sehr umfangreiche Sammlung statistischer Analysetools und lässt sich vergleichsweise einfach um selbst geschriebene Funktionen erweitern.

Mit Hilfe der bereits vorhandenen Analysetools lassen sich beispielsweise Muster innerhalb von Daten erkennen, Abhängigkeiten zwischen verschiedenen Merkmalen modellieren und Prozesse optimieren. Entsprechende Analysen könnten u. a. bei der Prognose von Umsätzen oder Ausschussware, der Modellierung von Kundenverhalten (inkl. Churn Prediction) oder im Bereich der Kundensegmentierung nützlich sein – selbstverständlich sind viele weitere Anwendungsfälle denkbar. Darüber hinaus ermöglicht R die Erstellung automatisierter Berichte sowie die Entwicklung eigenständiger und anwenderfreundlicher Analyse-Tools.

Im Rahmen der **Münsteraner R-Kurse** erhalten Sie die Gelegenheit, diese Konzepte auf wissenschaftlicher Basis und unter ständigem Einbezug der Praxis im Data Science-Kontext gewinnbringend für Ihre berufliche Tätigkeit nutzbar zu machen.

› DOZENTEN

Die Lehrenden der Münsteraner R-Kurse verantworten die inhaltliche Ausgestaltung und Weiterentwicklung dieser Seminare. Alle drei Beteiligten sind Lehrende des Instituts für Wirtschaftsinformatik der Westfälischen Wilhelms-Universität Münster:

- › Prof. Dr. Heike Trautmann
- › Dipl.-Inf. Jakob Bossek
- › Dr. Pascal Kerschke

› ANMELDUNG UND KONTAKT

Weitere Informationen zu den Münsteraner R-Kursen und zur Anmeldung finden Sie auf unserer Homepage unter www.wwu-weiterbildung.de/r-kurse

oder kontaktieren Sie uns unter:

WWU Weiterbildung gemeinnützige GmbH
Joshua Meyer
Königsstr. 47 | 48143 Münster

Tel.: 0251 83-27101
Fax: 0251 83-21709

E-Mail: j.meyer@uni-muenster.de

www.facebook.com/wwuweiterbildung

MÜNSTERANER **R**-KURSE UNIVERSITÄRE WEITERBILDUNGSSEMINARE IM DATA SCIENCE-KONTEXT





WWU WEITER BILDUNG



› INHALTE

BASISKURS

Der Basiskurs richtet sich an Teilnehmende, die die statistische Software R zur Analyse und Visualisierung eigener Daten erlernen möchten und bisher noch keine Vorkenntnisse besitzen oder zumindest eine Auffrischung ihrer Kenntnisse wünschen.

Neben der Installation und Bedienung wird vor allem die praktische Anwendung von R bei der Durchführung einfacher statistischer Analysen an realen Daten demonstriert und geübt. Im Anschluss daran werden einfache Techniken vermittelt, mittels derer sich R leichter und effizienter bedienen lässt, sowie einige grundlegende Methoden aus der Statistik vorgestellt, die zur Analyse eigener Daten verwendet werden können.

VERTIEFUNGSKURS Data Science und Machine Learning

In diesem Kurs für Teilnehmende mit R-Grundkenntnissen werden moderne statistische Verfahren zur Aufbereitung und Verarbeitung von Datensätzen sowie Methoden des maschinellen Lernens zur Analyse komplexer Klassifikations- und Regressionsprobleme vorgestellt.

Es werden vor allem nützliche Werkzeuge zur Modellierung maschineller Lernverfahren anhand praktischer Beispiele präsentiert. Dabei behandeln wir u. a. auch Themen wie die verbesserte Anpassung der Modelle an die Daten, eine korrekte Bewertung der Modellqualität sowie die richtige Interpretation der gefundenen Modelle und die Optimierung statistischer Lernverfahren.

Den Kursteilnehmenden werden verschiedene R-Funktionen präsentiert, die einen deutlich effizienteren Umgang mit komplexen Problemen bei großen Datenmengen ermöglichen.

VERTIEFUNGSKURS Effiziente Datenverarbeitung

In diesem Kurs, der R-Grundkenntnisse voraussetzt, lernen die Teilnehmenden, auch große Datenmengen mit R effizient verarbeiten zu können. Sogenannte *Split-Combine*-Ansätze sind generell geeignet, einen Datensatz in mehrere Datensätze aufzuspalten (*split*), Operationen auf den Teildatensätzen durchzuführen und die Ergebnisse wieder zusammenzuführen (*combine*).

Häufig ist die Split-Combine-Operation allein nicht ausreichend: Datenverarbeitung verlangt oftmals die Anwendung mehrerer hintereinander auszuführender Operationen. Der entsprechende Quelltext wird lang, unleserlich, fehleranfällig und die Übersicht leidet. *Pipelines* und *chaining* können oft helfen, die Übersicht zu wahren und Speicher einzusparen. Die Teilnehmenden lernen, diese Konzepte anzuwenden, um Alltagsprobleme der Datenverarbeitung mit R lösen zu können.

Kleine Datensätze stets im Arbeitsspeicher zu halten, ist unproblematisch. Große Datenmengen (Stichwort: *Big Data*) hingegen können zu massiven Performanceeinbrüchen führen. In diesen Situationen können Datenbanken helfen, große Mengen (un-)strukturierter Daten z. B. auf einer Festplatte zu speichern. R nutzt dann nur jene Daten, die im aktuellen Verarbeitungsschritt notwendig sind. Den Teilnehmenden wird in diesem Kurs entsprechend vermittelt, wie sie mithilfe von R effizient mit großen Datenmengen umgehen können und welche Bedeutung Datenbanken haben.

VERTIEFUNGSKURS Paket-Entwicklung in R

Einer der Gründe für die zunehmende Beliebtheit und den Erfolg der statistischen Programmiersprache R ist die Vielzahl von Erweiterungen in Form von *R-Paketen*, die über das *Comprehensive R Archive Network* (CRAN) oder *BioConductor* zur Verfügung stehen. In diesem Kurs für Teilnehmende mit soliden R-Kenntnissen wird die Entwicklung von Paketen thematisiert. Neben offensichtlichen Gründen, ein Paket zu entwickeln – wie beispielsweise Wiederverwendbarkeit und Reproduzierbarkeit von Forschungsergebnissen – sind R-Pakete darüber hinaus dazu geeignet, Code in Projekten zu strukturieren. Den Teilnehmenden wird dazu das nötige Handwerkszeug zur gelingenden Paket-Entwicklung in R an die Hand gegeben.

› TERMINE

- › **BASISKURS** 07.11.–10.11.2018 | 4 Tage | 1.700 €
- › **VERTIEFUNGSKURS Data Science und Machine Learning** 28.11.–01.12.2018 | 4 Tage | 1.700 €
- › **VERTIEFUNGSKURS Effiziente Datenverarbeitung** 16.01.–17.01.2019 | 2 Tage | 850 €
- › **VERTIEFUNGSKURS Paket-Entwicklung in R** 11.02.–13.02.2019 | 3 Tage | 1.275 €

Darin enthalten sind die Teilnahme an den Lehrveranstaltungen, die Seminarunterlagen sowie die Tagungsverpflegung. Diese umfasst die Tagungsgetränke, einen Mittagsimbiss und die Kaffeepausen. Die Seminare sind umsatzsteuerbefreit gemäß § 4 Nr. 21a (bb) UStG.