

# Datenanalyse mit R

Liebe Teilnehmerinnen und Teilnehmer,













in diesem vollständig digitalisierten Selbstlernkurs erwarten Sie praktische Lehrvideos (🎬), angeleitete Do-It-Yourself-Tutorials (📋), grundlegende Textbausteine (📄) und zielführende Musterlösungen (🔍). Die verwendeten Lehrmaterialien sind zu 100% mit den Inhalten meiner Präsenzveranstaltungen an der Universität zu Köln (ProfessionalCenter) sowie der RWTH Aachen (Lehrerbildungszentrum) als externer Lehrbeauftragter identisch und ermöglichen somit eine orts- und zeitunabhängige Wiederholung aller relevanten Inhalte. Einfach auf die entsprechenden Icons klicken und Sie gelangen direkt zu den Lehrmaterialien.

Ihnen viel Spaß beim Einstieg in die Datenanalyse mit R!


Dennis Klinkhammer

Icons  
anklicken  
und los  
geht es...

## (I) GRUNDLAGEN














Zeit	Inhalt	Link (YouTube)	Link (GitHub)
20 Min	Zum Start gibt es eine Einführung in die <b>wissenschaftlichen Gütekriterien</b> , welche auf einen <b>adäquaten Datenumgang</b> vorbereiten sollen.		
11 Min	Das erste Video ermöglicht den Einstieg in <b>R als Programmiersprache</b> und stellt die <b>Grundlagen</b> der grafischen Benutzeroberfläche von <b>RStudio</b> vor.		
08 Min	Der <b>TREES</b> Datensatzes ist ein erstes <b>Beispiel einer quantitativen Datenanalyse</b> , ohne dass statistische Vorkenntnisse erforderlich sind.		
30 Min	In dieser <b>Übungsaufgabe</b> werden <b>erste Befehle</b> zum <b>Umgang mit Daten</b> und zur <b>Auswahl von relevanten Fällen</b> vorgestellt.		
05 Min	Komplexere quantitative Datenanalysen erfordern sogenannte <b>Analysemodelle</b> , welche in R bspw. mit <b>DiagrammeR</b> angelegt werden können.		
12 Min	Der <b>SWISS</b> Datensatz verdeutlicht die <b>Bedeutung theoretisch fundierter Analysemodelle</b> hinsichtlich der Interpretation von statistischen Befunden.		
30 Min	Die zweite Übungsaufgabe widmet sich dem <b>MTCARS</b> Datensatz, in dem <b>die Analyseschritte aus dem vorherigen Video wiederholt</b> werden können.		
20 Min	Die zu berücksichtigenden <b>Herausforderungen bei der Datengewinnung</b> sollen die <b>Komplexität bei der weiteren Datenanalyse</b> verdeutlichen.		

## (II) FORMELSAMMLUNG









Zeit	Inhalt	Link (YouTube)	Link (GitHub)
05Min	Eine <b>praktische Übersicht</b> über die in den nachfolgenden Lehrvideos und Do-It-Yourself-Tutorials ausgewiesenen <b>statistischen Formeln</b> .		

# Datenanalyse mit R

## (III) DATENANALYSE

Zeit	Inhalt	Link (YouTube)	Link (GitHub)
20 Min	Für ein besseres Verständnis der statistischen Formeln empfiehlt sich vorab ein Blick auf die unterschiedlichen <b>Skalenniveaus</b> .		
12 Min	In diesem Video werden im Rahmen der <b>univariaten Statistik</b> die <b>Lagemaße</b> und die <b>Streuungsmaße</b> vorgestellt.		
20 Min	Die Standardabweichung ermöglicht die sogenannte <b>z-Transformation</b> zur Überführung von <b>Rohwerten</b> in vergleichbare <b>Normwerte</b> .		
19 Min	Die <b>bivariate Statistik</b> fokussiert über <b>Korrelationen</b> , <b>Chi-Quadrat-Tests</b> und <b>t-Tests</b> den Zusammenhang zwischen jeweils zwei Variablen.		
30 Min	Der <b>Chi-Quadrat-Test</b> lässt sich <b>von Hand</b> ausrechnen, wofür ein passendes Beispiel im Rahmen der dritten <b>Übungsaufgabe</b> bereitsteht.		
30 Min	Zur Wiederholung des <b>t-Tests</b> wird in dieser <b>Übungsaufgabe</b> wieder auf R und den <b>ToothGrowth</b> Datensatz zurückgegriffen.		
14 Min	Ausgehend von der <b>linearen Regression</b> wird im Rahmen der <b>multivariaten Statistik</b> ebenfalls die <b>logistische Regression</b> vorgestellt.		
18 Min	Die <b>Faktorenanalyse</b> und die <b>Clusteranalyse</b> zählen zu den <b>komplexitätsreduzierenden Verfahren</b> und fassen Variablen bzw. Fälle strukturiert zusammen.		
30 Min	Die fünfte <b>Übungsaufgabe</b> greift auf den <b>BFI</b> Datensatz zurück, um die Befunde der Korrelation mit denen der <b>Faktorenanalyse</b> abzugleichen.		

## (IV) MACHINE LEARNING

Zeit	Inhalt	Link (YouTube)	Link (GitHub)
20 Min	In der <b>theoretischen Einführung</b> zum Machine Learning werden das <b>Supervised Machine Learning</b> und <b>Unsupervised Machine Learning</b> vorgestellt.		
06 Min	<b>Praktische Einführung</b> zur Erstellung von Algorithmen sowie <b>Trainings- und Validierungsdatensätzen</b> am Beispiel des <b>TREES</b> Datensatzes.		
12 Min	Analyse des <b>IRIS</b> Datensatzes mittels Machine Learning Algorithmen aus dem Classification and Regression Training – <b>Caret Package</b> .		
30 Min	Die letzte <b>Übungsaufgabe</b> wiederholt die Schritte aus dem Video und ermöglicht einen <b>Abgleich mit den Ergebnissen aus der ersten Übungsaufgabe</b> .		
10 Min	<b>Resampling</b> präzisiert nicht nur einzelne Parameter im Rahmen der Statistik, sondern ist auch Grundlage für ein <b>belastbares Machine Learning</b> .		
20 Min	Übersicht über gängige <b>Herausforderungen in der Anwendung</b> von Machine Learning Algorithmen und <b>Handlungsempfehlungen</b> für die Praxis.		