JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

GGS
Gießener Graduiertenzentrum
Sozial-, Wirtschafts- und Rechtswissenschaften

Workshop

Advanced Analytics with R - Causal Inference

Organizational details

Instructor:	Dr Tobias Keller
Date & Time:	October 6, 7 & 8, 2022, 9.00 am – 5.00 pm
Location:	Room 002, Licher Str. 68, 35394 Giessen
ECTS:	3
Max. participants:	12

Objectives

Refresh your basic R programming knowledge and add important techniques for data manipulation, automation, parallelization, debugging, and reproducible research. Learn when and how to apply methods for causal inference such as Propensity Score matching and Difference-in-Difference (DID) Designs as well as how to implement these techniques using R.

Content and methods

Structure:

- Recap of required knowledge (approximately 1 to 2 hours)
    - Basic R programming and using RStudio
    - Running and interpreting linear regression models
- Advanced R programming techniques (approximately 1 to 1.5 days)
    - Advanced data manipulation
    - Automation and parallelization
    - Debugging
    - Reproducible research
- Introduction to causal Inference with R (approximately 1 to 2 days)
    - Fundamental problem of causal inference
    - Propensity score matching
    - Difference-in-difference design
    - A glimpse of more methods for causal inference

The approximated time we will spend on the individual chapters is indicated above. We may deviate from this schedule depending on the pace and on questions asked during the course.

Methods:

The course consists of lectures to build the theoretical background as well as hands-on tutorials and exercises using R. Participants will learn by examples and exercises from the instructor's experience in research and practice.

Hands-on tutorials and exercises make up about 50% of the course time. The exercises will be based on exemplary datasets that will be provided to the participants before the course.

Target group

Doctoral candidates or postdoctoral researchers who are already conducting empirical research using R or intend to do so. It is required that the participants have basic R programming knowledge. In particular, the participants should be comfortable with performing the following tasks using R:

- Data transformation: filtering, selecting columns, calculating new variables
- Plotting: histograms, scatter plots including regression lines
- Linear regression using the function lm() including factors and interactions

Participants will be expected to bring their own laptop computers with the following software installed:

To gain the ECTS credit points participants will have to:

- Refresh their prior knowledge of R programming
- Install R, version 4.0 or higher, downloadable here: https://cran.r-project.org/
- Install R Studio 2021.09.0 or higher (free Open Source Desktop edition), downloadable here: https://www.rstudio.com/products/rstudio/download/#download
- Install the following packages:
  - install.packages(c("tidyverse","data.table","foreach","MatchIt","rdrobust","rdrobust","rdlocrand","rddensity","rdmulti","rdpower","locpol"))
- Ensure a working internet connection so that packages can be downloaded as required during the course

Course language

English (German, if only German participants)

Please note: As this is not an English language course proficiency in English at the C1 level of competency is required.

About the Instructor

Dr Tobias Keller has been working as data scientist at ZERO.ONE.DATA, the big data start-up of Deutsche Bahn AG since 2016. He consults on and applies machine learning and statistics for artificial intelligence systems in a big data environment. Furthermore, he teaches data science at Deutsche Bahn, in the doctoral education programmes at Justus Liebig University Giessen, and in the doctoral programme and master programme at WHU – Otto Beisheim School of Management. His research interests include machine learning and artificial intelligence, finance and accounting, strategic management.

Registration

By **September 26, 2022** via e-mail at info@ggs.uni-giessen.de.