



IMPRS-PHDS elective course

ThinkData: Introduction to Administrative Register Data

Christian Dudel

Start: 17 May 2023 End: 17 May 2023

Location: Max Planck Institute for Demographic Research (hybrid)

Instructor:

Christian Dudel

Course description

Administrative register data comes with many promises: rich and large data sets covering complete populations, and without issues of unit- and item-non-response. Because of this, register data is increasingly used in the social sciences. This is particularly true for studies covering the Nordic countries, for which register data has become the de-facto standard for many research topics. Other countries are in the process of catching up, and more and more register data sets are becoming available to researchers.

In practice, using administrative register data can be very challenging. The largeness of register data is a double-edged sword, making data processing and computations slow. Moreover, register data is complex, especially when several registers have to be merged. Because of this the data preparation process is error prone. Also, registers are not structured to meet the needs of researchers but with administrative processes in mind, adding another layer of complexity.

This course provides an introduction to administrative register data, aimed at researchers interested in starting to use such data. We will work on an example which covers some of the challenges of register data, and complex longitudinal data more generally, and we will see how these challenges can be tackled. This includes handling of big data, merging of files in different formats, and consistency checks. We will use R, but most of the ideas and concepts also apply to other statistical software.

Organization

The course will be held at the Max Planck Institute for Demographic Research. Students can attend on-site or online (hybrid format). This one-day course requires students to work full-time on 17 May, the day of the course. The course ends at 5 pm local time (CEST).

Course prerequisites

Fluency in R is desirable but not a requirement. Participants are expected to bring their own laptops.

Examination

There is no examination in this course. A certificate of attendance will be provided to all participants.

Tuition

There is no tuition fee for this course.

General readings

Roxanne Connelly et al. (2016): The role of administrative data in the big data revolution in social science research. *Social Science Research* 59, 1-12. https://doi.org/10.1016/j.ssresearch.2016.04.015

Course materials will be available here: https://github.com/christiandudel/ThinkData2023

Admission

Course attendance is limited to doctoral students enrolled at IMPRS-PHDS and MPIDR research staff.

How to apply

- IMPRS-PHDS students are asked to register for this course by ticking the appropriate box on the PHDS Graduate Workshop registration survey or by emailing phds@demogr.mpg.de.
- MPIDR scientific staff are kindly requested to register for this course by sending an email to phds@demogr.mpg.de.
- Application deadline is 9 May 2023.

Questions about the course should be sent to phds@demogr.mpg.de.