STV2020 – Social Science data analysis and programming

Course content

The course offers an introduction to programming-based problem solving for social scientists. With ongoing digitalization in the public sector and automation of services, digital competence is in demand among employers in both the public and private sectors.

The course introduces a number of problems and solutions in social science data processing with applications in R. We start with a number of general programming topics, followed by efficient processing of different data structures and how data can be combined using SQL and Tidyverse. Secondly, we look at special challenges related to space and time. The spatial dimension introduces GIS techniques. Towards the end, we see how machine text analysis can be used to automate data collection and we look at how we can effectively visualize different types of data.

The course provides a good basis for independent work with social science information.

Learning Outcome

Having concluded this class students will:

**Knowledge**

* Efficiently process different types of data
* Be able to write R code
* Master the whole online process from data collection via analysis to visual presentation.
* Be able to define a research problem that can be answered by information available online

**Skills**

* be familiar with programming in R, and use data structures, write loops and more efficient options, do condition tests and write your own functions
* be able to write R-code that retrieves data from web pages, analyze these and present the results using tables and figures
* be able to make an interactive presentation of your research results

**General competence**

* know how to collect, prepare, analyze and present relevant data to answer social science questions

Teaching

Lectures and seminars with compulsory activities

Exam

* term-paper

Recommended background knowledge

stv1020