

# Computer programming E140

Christian Dudel

September XX, 2020

# What will be covered in this course?

- ▶ Software: R and RStudio
- ▶ Mostly basic things in these programs
- ▶ Course will not cover many things
- ▶ Course will not be deep

# Goals

At the end of this course...

- ▶ ...you understand basic concepts of R
- ▶ ...you can do a basic analysis in R
- ▶ ...you know where you can learn more about R

# Prerequisites

- ▶ Basic demographic knowledge (e.g., you know what a 'rate' is)
- ▶ Basic statistical knowledge (e.g., you know what a 'mean' is)
- ▶ First experience using statistical software (Stata, Excel, SAS, SPSS, R, ...)

## Already an expert?

- ▶ Write a function that takes any integer as input and returns TRUE if the integer is a prime number, otherwise it returns FALSE
- ▶ Write a function that takes any integer as input and returns its amicable number if it exists, otherwise it returns FALSE
- ▶ Check one of the many tasks on [rosettacode.org](http://rosettacode.org) solved with R
- ▶ Replicate one of the many replicable articles available at Demographic Research
- ▶ Work on your own analysis

# Contact

- ▶ Email: [dudel@demogr.mpg.de](mailto:dudel@demogr.mpg.de)
- ▶ Office: 359 (3rd floor, east wing)
- ▶ Twitter: @c\_dudel
- ▶ Website: <https://sites.google.com/view/cdudel>

# Topics I work on/I am interested in

- ▶ Labor markets and aging
- ▶ Pensions and inequality
- ▶ Fertility, in particular fertility of men
- ▶ Longitudinal data analysis, in particular multistate modeling
- ▶ Identification and causal inference
- ▶ Survey methodology (variance estimation, missing data)

# Course schedule

1. session: Date and time not set yet
2. session: Date and time not set yet
3. session: Date and time not set yet
4. session: Date and time not set yet
5. session: Date and time not set yet



# Materials

Materials will be available from GitHub, also mirrored on OSF:

- ▶ <https://github.com/christiandudel/EDSD2020>
- ▶ <https://osf.io/c6jru/>

Materials will mostly consist of R code and some slides.

# Assignment: Overview

- ▶ One assignment
- ▶ Assignment handed out on XX.YY.
- ▶ Deadline: ZZ.ZZ.
- ▶ Assignment will consist of several tasks: “Do this and that with R”, “Solve this problem with R”, etc.

# Assignment: Your solutions

- ▶ You submit R code as solutions
- ▶ R code should be commented, explaining what is happening
- ▶ Code should work “out-of-the-box” without errors

# Assignment: Deadline

- ▶ Deadline assignment: XX.ZZ.YY, 12am
- ▶ Send your solutions to me (dudel@demogr.mpg.de)
- ▶ You will get a confirmation (might take a few days, sorry)
- ▶ I might get back to you if I have problems with your file(s)
- ▶ It is your responsibility that your files are working!

# Assignment: Groups

- ▶ You can work in groups
- ▶ Actually, I strongly suggest you work in groups!
- ▶ Please not more than five people per group
- ▶ Please submit your solutions only once per group
- ▶ Make clear who is member of the group when submitting

## Assignment: Grading

- ▶ You can either “pass” or “fail”
- ▶ Your code should... –...work “out-of-the-box” –...be well-documented: Comments! –...should be (somewhat) efficient. If one step can do the work then don't use two or more!

# Assignment: Summary

- ▶ One assignment consisting of coding tasks
- ▶ You submit code as solutions
- ▶ You can work in groups
- ▶ Pass/fail

# What is R?

- ▶ R is an open source statistical programming language
- ▶ First release in 1995
- ▶ Used for data analysis and statistical programming



# Why use R?

- ▶ Free, open source
- ▶ Can easily be extended
- ▶ Around 15,000 packages available
- ▶ De facto standard in statistics, commonly used in both science and industry
- ▶ Tons of R-related things: Books, journals, forums, conferences, ...
- ▶ Many methods are already implemented in R

# Why use RStudio?

- ▶ R is the programming language
- ▶ RStudio is a tool to use R more efficiently
- ▶ Features: – Syntax highlighting, code folding – Project management (e.g., GitHub) – Markdown support – ...

# Disclaimer

- ▶ R is not the only statistical software and it is fine if you prefer something else
- ▶ RStudio is not the only IDE/editor for R (ESS, RKWward, Tinn-R, ...)
- ▶ R can be used in many different ways
- ▶ Example: base R vs tidyverse vs data.table vs specialized packages
- ▶ I do things in certain ways, and this course will follow that
- ▶ This does not mean that the solutions from this course are the only or the best way to do things

# What do you need to get started?

- ▶ R: <https://cran.r-project.org/>
- ▶ R-Studio: <https://www.rstudio.com/>