IMPRS

INTERNATIONAL MAX PLANCK
RESEARCH SCHOOL
population. health. data science.

MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

IDEM 181

Data visualization – the art/skill cocktail

Start: 14 June 2021
End: 18 June 2021

Instructor: Ilya Kashnitsky, University of Southern Denmark

Location: Online course. Link tba.

Course description

In the ever-growing universe of dry academic texts, impressive and efficient graphics are quite rare. Driven by widespread software legacy issues and mostly outdated limitations imposed by traditional scientific publishers, researchers often consider producing high quality graphics as a peripheral optional task – "if time allows" (spoiler: it won't). Yet, modern tools place data visualization in the focus of research workflows when it comes to conveying the results. Hence, the ability to turn a large dataset into an insightful visualization is an increasingly valuable skill in academia.

The course aims to empower the participants with the flexibility that the R+tidyverse framework gives to visualize data (the practical examples mostly use demographic data). The course covers some aspects of data visualization theory and best/worst practice examples, but it's primarily practice oriented including live coding sessions and short lecture/showcase parts.

Practical coding sessions start from basic introduction to tidy data manipulation and ggplot2 basics. Next, practical examples cover the creation of certain most useful types of plots. Important data visualization choices and caveats are discussed along the way. Special attention is devoted to producing geographical maps, which are no longer the luxury of professional cartographers but have turned, with the help of R, into yet another data visualization type. Going beyond ggplot2, the course presents an introduction to interactive data visualization.

Organization:

Each of the five course days will consist of one two-hour lecture 14:00-16:00 CEST (Central European Summer Time), followed by a one hour discussion time (16:00 – 17:00 CEST), which is optional for the participants.

Monday, June 14: BASICS
- Basic dataviz principles
- Impressive dataviz showcases
- Tidy approach to data
- {ggplot2} basics

Tuesday, June 15: TUNE-UP
- More advanced {ggplot2}
- Colors in dataviz
- Themes and fonts
- Population pyramids and animation

Wednesday, June 16: TOOLBOX
- Useful types of dataviz
- Dotplots – the most neglected and powerful type of dataviz
- Heatmaps, equality-line, ggridges, treemap

Thursday, June 17: MAPS

- The basics of map projections
- {sf} – the game changer in #rspatial, `geom_sf`
- Useful spatial processing tricks

Friday, June 18: ROCK

- Creative legends: biscale, tricolore
- Interactivity: plotly, gganimate
- Intro to shiny

Course prerequisites

Participants should have basic experience in using R. For those starting from scratch, it's a good idea to take some of the online introductory courses (swirl R package https://swirlstats.com is one nice option; RStudio provides brilliant Primers https://rstudio.cloud/learn/primers). Participants need a laptop or desktop computer with the latest versions of R and RStudio installed. More information regarding the R packages to install will be sent before the course starts.

Evaluation

Four small assignments will be provided during the course.

General readings

I suggest two recent books that are both freely available online
– by Claus Wilke – https://serialmentor.com/dataviz
– by Kieran Healy – https://socviz.co

HOW TO APPLY

Please apply via https://www.demogr.mpg.de/go/idem181 and include as ***single pdf file***
(max. 10 MB), **in English**:

(1) A curriculum vitae, including a list of your scholarly publications (max. 2 pages).
(2) A one-page statement of your research and how it relates to the course. Please include a short description of your knowledge of data visualization and your fluency in R.
- Application deadline is **8 May 2021**.
- Applicants will be informed of their acceptance by 21 May 2021.

Tuition

There is no tuition fee for this course.

Recruitment of students

- Applicants should either be enrolled in a PhD program or have received their PhD.
- The selection will be made by the MPIDR based on the applicants' scientific qualifications.