

# Spatial data analysis with R

CSC – 1.-3. Nov 2021

Marko Kallio – [marko.k.kallio@aalto.fi](mailto:marko.k.kallio@aalto.fi)

# Introduction round

- Let us know
  - Who you are?
  - Which institution you're coming from?
  - Do you have experience with R?
  - Do you have experience with spatial data in R?

# Introduction

- Doctoral researcher at Aalto University, School of Engineering
  - Themes include:
    - Methodological research
    - hydrology, environmental modelling
    - Spatial interpolation
    - Spatio-statistical methods in downscaling
  - Master's degree in Geoinformatics
  - Bachelor's degree in Environmental Engineering (Water management)
- Working daily with R
  - Small scripts to support fellow researchers
  - Package development: hydrostreamer, (dasymetric), (Smoodjustment)

# Contents

- Day 1
  - Vector data in R, vector data manipulation and queries
- Day 2
  - Spatial analysis with vector data – spatial autocorrelation, spatial clustering, and spatial models
- Day 3
  - Raster data in R, raster data manipulation, map algebra and spatial modelling with rasters

# Learning outcomes

- You'll have an idea how to get started with spatial data in R
- You can switch to programmatic workflows from desktop QGIS (at least for some operations)
- Most of all, I hope you'll have learnt something new not only about R, but also from the methods we will be using. Even if you're already experienced.

# What this is not include:

- R basics. I assume you already know how R works.
- Comprehensive look at the R spatial ecosystem. There are far too many spatial packages in R to go through. This only skims the surface.
- Visualisation course – we'll not learn how to make publication ready figures. Although we go through some basic visualisation, these are mainly how I personally use them when working interactively with spatial data. For proper maps, I use QGIS + Affinity, but there are many good packages for visualisation of spatial data in R.

# Beyond this course

- CRAN Task view of Spatial Analysis  
<https://cran.r-project.org/web/views/Spatial.html>
- Geocomputation with R  
<https://geocompr.robinlovelace.net/index.html>
- Spatial Data Science with R  
<https://rspatial.org>
- And so many more...

# Getting started

- All of the material and data we'll be using are hosted at <https://github.com/csc-training/r-spatial-course>
- Head to notebooks.csc.fi and log in with CSC or HAKA credentials

- Choose Rstudio environment for DAKI and teaching activities.

Rstudio environment for DAKI and teaching activities

This Rstudio environment is based on the latest version of R (v 4.1.0) along with the necessary R packages for teaching as well as DAKI project activities. This environment is short-lived and all data is destroyed at the end of the session.

Lifetime: 10h

➤ Launch new

**Click “launch new” ONLY ONCE!** Be patient. Click open in browser when it is ready.

Rstudio environment for DAKI and teaching activities

This Rstudio environment is based on the latest version of R (v 4.1.0) along with the necessary R packages for teaching as well as DAKI project activities. This environment is short-lived and all data is destroyed at the end of the session.

Lifetime: 10h

State	Name	Time Left	Access	Actions	Details
✓	pb-drew-the-elegant	9h	Open in browser	✕ Destroy	Details

➤ Launch new

?



# Getting started

- Copy the password given in prompt.
- Login using
  - Username: rstudio
  - Password: paste the password you just copied.
- Rstudio is now running on Puhti, and you'll be able to use it through your browser!
- One last thing: you'll need to clone the Github repository to get the material and data to the Rstudio session. →

# Getting started

- Click on the terminal

- In the terminal window, type  
*git clone https://github.com/csc-training/r-spatial-course.git*  
and press enter

Make sure you have everything: in the *files*-pane, check that you have a new folder called *r-spatial-course*.

The materials can be found under folders *Day 1, Day 2, Day 3*

