Notes on Ch1 - Introduction

Norman LIm

2025-06-28

What is geocomputation?

- activities, such as academic research, software developement, and practical applications, that use geographic data to solve problems, with a focus on reproducibility, flexibility, and tool development.
- first used in 1996 during the first conference on the subject
- closely related to Geographic Information Science (GIS), Geomatics, Geoinformatics, Spatial Information Science, Geoinformation Engineering, and Spatial Data Science.
- Here, "science" means reproducible and falsifiable

Why use open-source software for geocomputation?

Historically: - Early geographers used barometers, compasses, sextants, etc. to learn about the world - With the invention of the marine chronometer in 1761, it became possible to calculate longitude at sea - but there was a big shortage of data and tools for geographic analysis

Now: - there is no such shortage of tools and data - most phones are equipped with GPS - we can use satellites and semi-autonomous vehicles - we have instruments that generate gigabytes of data daily - data can be accessed through various data storage media and APIs

Therefore: - we need to use tools that are future-proof (because of the rapidly changing hardware, software, and data landscapes) - we need tools that can keep up with this rapid rate of development - we need tools that promote reproducibility (because being able to replicate findings is vital; with open-source software, "anyone" can check your findings) - we need tools and software that promote collaboration the creation of community where one can get support/feedback far quicker than the support team of a proprietary product

This is a good analogy: Proprietary products are like "monolithic empires" that are difficult to maintain, while open-source products are like a "federation" of modular tools that can be combined in different ways.

Why R for geocomputation?

- R is open-source and is used for statistical computing and graphics
- RStudio is a good IDE for geocomputation and data visualization
- R is an object-oriented and functional programming language (as per Wickham)
- R includes may bridtes to plenty of GIS softare and geolibraries, and functions
- Creation of new tools/libraries is easier (compared to lower-level languages) like C or FORTRAN
- R has libraries that facilitate access to other programming languages (like C++ and Python)
- R can create interactive maps

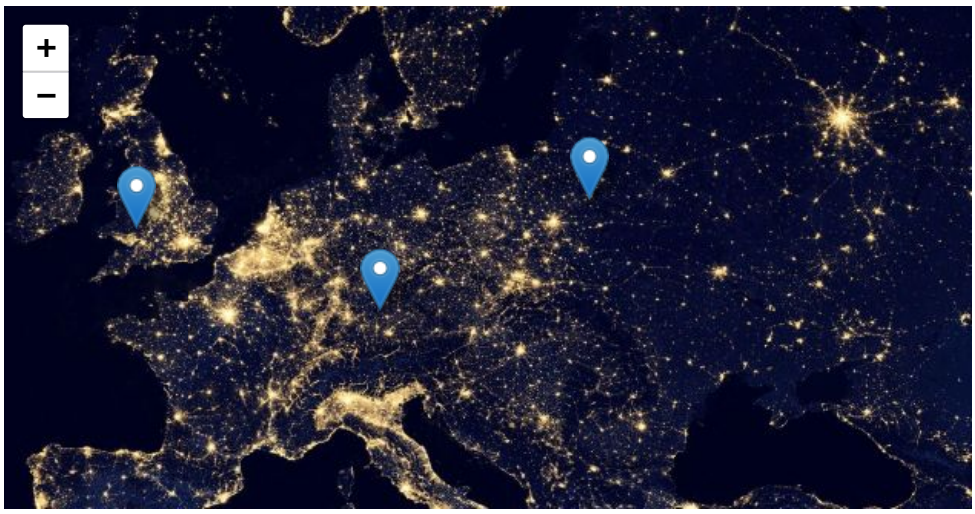
Example showing how "easy" it is to make interactive maps in R:

```r
library(leaflet)

popup = c("Robin", "Jakub", "Jannes")
leaflet() |>
  addProviderTiles("NASAGIBS.ViirsEarthAtNight2012") |>
  addMarkers(
    lng = c(-3, 23, 11),
    lat = c(52, 53, 49),
    popup = popup
  )
```

```
## PhantomJS not found. You can install it with webshot::install_phantomjs(). If it is installed, please

## file:////tmp/RtmpfnvMMc/file23c1c6a38b8c6/widget23c1c50f73bc0.html screenshot completed
```



- A few years ago, it would have been difficult to produce the figure above using R or any open-source language.

- development of libraries such as **knitr** and **leaflet** made this possible

Software for geocomputation

- R and Python: both are interpreted language
- C++ and Java: needs compilation, but usually runs faster once they have been compiled
- QGIS, GRASS GIS, SAGA: GIS software

R or Python?
- it doesn't really matter
- Python is a good general-purpose program
- Many geoalgorithms can be accessed from the command line of Python
- For spatial statistics and predictive modeling, R is second-to-none

R's spatial ecosystem

- the package `sf` (built from `sp`)
- terra
- r5r
- stars
- leafem
- spdep
- tmap

History or R-spatial

- R's spatial capabilities originated from early spatial packages in the S language
- `spatial`, `sgeostat`, and `splancs` from year 2000
- `spatsat` from 2001
- `sdep`, `maptools` from 2017
- `rdal` and `sp` for reading shapefile file format
- `GDAL` and `PROJ`: high-performance
- `rgdal` from 1993 provided GDAL bindings for R
- `rgeos` from 2010, developed during Google Summer of Code in 2010
- `raster` from 2010
- `terra` which superseded `raster` (both terra and ruster can work with datasets that are too large to fit on RAM)
- `spgrass6`, `rgrass7`, `rgrass`, `qgisprocess`, `SAGA`, bridge software between GIS softare and RStudio
- `RgoogleMaps` overlays spatial data on top of the `basemap`
- `ggmap`: like ggplot, but for geospatial data