Working with Spatial Data: Applications in R

SYLLABUS

Given by: Arijit Ghosh (arijit.ghosh@rwi-essen.de)

The course deals with specifics of applications in spatial data, nuances which all economics graduate students can make good use of. In this seminar, we go through the most commonly used geographic information systems (GIS) tools and follow a simple framework of learning how to work with spatial data. Within this framework, there are three main learning objectives:

- Visualizing maps
- Finding distances between objects
- Extracting values at locations

Using just the learning outcomes above, it is possible to understand the depth of spatially explicit economic research spanning applied microeconomics and applied macroeconomics. More importantly, it can help with building sophisticated research ideas.

For this course, in-class exercises requiring the use of RStudio (an IDE for the R programming language) have been designed. R provides ample flexibility in this regard: it contains a comprehensive set of accessible commands for use on geocoded data. Moreover, its learning curve is not steep.

The following describes the scope of the course in detail (all of the short course components below will be supplemented with in-class exercises):

- Brief introduction to the wide variety of spatial data and their sources
- Introduction to the spatial data packages in R
- Introduction to different types of spatial data such as rasters (for eg., satellite imagery on climate data, grid data on HDI, the RWI-GEO-GRID, etc.), shapefiles (for eg, the shape of a city or a road network), and points-data
- Reading and visualizing spatial data
- Overlaying different types of spatial data for analyses and visualizations
- Finding distances between lists of points (distances between wind-turbines, for example)
- Finding distances between list of points and list of polygons (distances between homes and national parks, for example)

- Extracting values from rasters at points (for eg., finding the slope at a given latitude-longitude coordinate)
- Extracting and aggregating values from rasters on polygons (for eg., finding the average population density in a radius of 1 km from a given location)
- A note on computation speeds for extraction algorithms

Prerequisites for this course: there are no prerequisites for this course. The course is designed keeping in mind the problems graduate students in economics might encounter while working with spatial data. Familiarity with R and RStudio is not necessary.

The course takes place on May 19 (Thu) and May 20 (Fri) with three 1.5 hour sessions on Thursday (09:15 - 10:45, 11:00 - 12:30, 14:00 - 15:30) and two 1.5 hour sessions on Friday (09:15 - 10:45, 11:00 - 12:30). It will take place in-person at the RWI. Having personal computers during this course would be ideal but RGS students may use their work desktops too. Hybrid/online formats will be considered in case Covid-19 developments take a turn for the worse by then.

Some motivating references (also genuinely fun reads):

Recent publications:

Harari, M. (2020). Cities in bad shape: Urban geometry in India. American Economic Review, 110(8), 2377-2421.

Fouka, V. (2020). Backlash: The unintended effects of language prohibition in US schools after World War I. The Review of Economic Studies, 87(1), 204-239.

Flückiger, M., Hornung, E., Larch, M., Ludwig, M., & Mees, A. (2022). Roman transport network connectivity and economic integration. The Review of Economic Studies, 89(2), 774-810.

JMPs from 2021/2022: Chopra, F. (2021). Media Persuasion and Consumption: Evidence from the Dave Ramsey Show. Available at SSRN.

Morales-Arilla, J. (2021). Autocrats in crisis mode: Strategic favoritism during economic shocks.

Levy, A. (2021). Housing Policy with Home-Biased Landlords: Evidence from French Rental Markets.

Seminal piece: Dell, M. (2010). The persistent effects of Peru's mining mita. Econometrica, 78(6), 1863-1903.