

# In-class lab 2

## Geospatial analysis in R KDI School Fall 2024

You can find all of the data in the [assignments/in-class2/data](#) folder. The datasets are:

- `meshblocks.shp`: a shapefile of administrative areas of New South Wales (NSW), Australia (specifically, just around Sydney)
- `NSW.shp`: a shapefile of the state of New South Wales (at a lower level of aggregation)
- `pollution.csv`: a dataset of pollution monitoring stations for seven days in NSW
- `pop.tif`: a raster with population values from WorldPop for the area around Sydney

As a quick note: you can still get full credit even if you do not get through all of these tasks. I would rather you do a few things well than many things poorly. While time pressures are an important consideration in the real world, it usually isn't the most important consideration. Quality is much more important. I will grade you on both the quality and the quantity of your work, but with a much larger emphasis on the former.

By the end of class, you need to turn in a raw markdown file AND the pdf output. Please make sure ALL STUDENTS' names are on the assignment. Here are your tasks for today:

- Let's start with the pollution data. Please load it and do the following:
  - Create a single observation for each station, with average pollution values for all seven days.
  - Turn the data into a spatial object.
- Now create a simple map. Please load the `NSW.shp` file and plot it, *along with the locations of the pollution monitoring stations*.
- We eventually want to extract pollution values for mesh blocks, but we need to do some cleaning first. I want you to create a raster for the entire state of NSW with interpolating pollution values. Please create two separate rasters:
  - One using simple inverse distance weighting
  - One using kriging
  - After you create the rasters, please plot them side-by-side on the same plot.

- Note: The resolutions of the rasters are up to you. However, remember that we are going to eventually use it to extract values for mesh blocks, so the raster cells cannot be too large.
- Now let's move on to the mesh blocks. Please load the `meshblocks.shp` file and do the following:
  - Extract the values of pollution for each mesh block.
  - Extract the values of population (sums) for each mesh block.
  - Plot both side-by-side on the same plot.
- The last step is not geospatial, per se. I'd like you to create a descriptive figure that shows the relationship between population counts and pollution at the mesh block level. For example, this could be a scatter plot with population on the x-axis and pollution on the y-axis.

Please treat this like a paper. Include a short description of figures and write in full sentences.