# R Shiny dashboard development exercise

## Goal

Your task is to create an R Shiny dashboard which follows the outlined design specifications below.

## Description

A huge part of the work our group does, is to simulate the malaria transmission on the country level for various settings and malaria intervention combinations. The term setting refers to geographical places such as provinces, districts, etc. which are varying in their characteristics (mosquito species, climate, present level of malaria (EIR, Entomological Inoculation Rate), ...).

You have been given a simplified sample dataset containing the output of such an epidemiological simulation for the hypothetical development of malaria in Switzerland. This is a fictional scenario and the dataset includes made up data. For the exercise, you should create a dashboard which visualizes the data. You are free to choose the layout, design, number of panels, tabs, etc. as well as the interactivity (e.g. mouse-overs, popups) to your liking.

The dashboard should include the following features:

1. UI elements allowing the user to choose the population age group
2. Visualization of the incidence and prevalence rate over time as well their reduction compared to the counterfactual/baseline.
3. Map(s) showing which intervention combination was deployed in each administrative unit (use **only** the NSP scenario).

Your target audience are people from the National Malaria Control Program (NMCP) which is responsible for planning the malaria intervention campaigns. Your dashboard should convince them of the feasibility of the planned interventions.

The data you have received is a small subset of real-world data which is commonly produced and used in our group. You have only one counterfactual (BAU, Business as usual) and one National Strategic Plan (NSP) in your data along with the two corresponding scenarios (prefixed with either „BAU“ or „NSP“). Real world data is often in the range of 1 million rows and up and thus, performance matters. Pretend that you are working with a large dataset and use any technique that you think would help with performance. Write the code as you would for a real-world project.

Please complete the following tasks in order to complete the exercise:

* ***Briefly*** summarize your approach from conception to development and deployment in a separate text document.
* State any assumptions or decisions you made along with why they were made (e.g. coding standards, library choices).
  + This can be done within the code.
* To ensure fairness among all candidates, we will not answer questions you may have while doing this exercise. In case there are things that are unclear to you, or additional assumptions you make and would like to explain, please feel free to write them as a separate document.
* **State if you used LLMs/AI code assistants and what you used them for.**
* Submit the above together with the dashboard code to [roland.goers@swisstph.ch](mailto:roland.goers@swisstph.ch) and [gianpaolo.pontiggia@swisstph.ch](mailto:gianpaolo.pontiggia@swisstph.ch). If you changed the underlying data, ensure that you submit it as well.

Notes

|  |  |  |
| --- | --- | --- |
| Column name | Type | Description |
| country | char | Country name (admin 0). |
| admin\_1...5+ | char | Each column contains the name of the administrative division. See [here](https://en.wikipedia.org/wiki/List_of_administrative_divisions_by_country) or [here](https://sedac.ciesin.columbia.edu/povmap/ds_defs_admin.jsp) for details. |
| scenario\_name | char | Unique identifier of the simulated scenario. |
| seed | int | Seed number. |
| EIR | num | EIR value. |
| year | int | Year of simulation outputs. |
| nHost | num | Population in given age group. |
| deployed\_int\_XYZ | bool | If intervention XYZ has been deployed in the given year. XYZ is the intervention name. |
| prevalenceRate | num | Prevalence rate. |
| incidenceRate | num | Incidence rate. |
| nUncomp | num | Number of uncomplicated cases. |
| tUncomp | num | Number of treated/reported uncomplicated cases. |
| nSevere | num | Number of severe cases. |
| tSevere | num | Number of treated/reported severe cases. |
| expectedDirectDeaths | num | Number of deaths. |

Table 1: Data structure.

* You can merge the data from the shapefile and the simulation data via the NAME and the admin\_1 column.
* You can choose third-party packages as you see fit. Some recommendations:
* sf : For handling geospatial data. You can load the shapefiles like so st\_read("path/to/shapefiles/") |> st\_transform(4326)
* leaflet for the creation of maps