**Assignment**

**Bayesian Statistics in Epidemiology**

**Important:**

Full marks are assigned based on the appropriate application and interpretation of the Bayesian workflow steps, **not** based on finding a “positive” association between variables.

For all assignments you can use the code and data in the Github repo (https://github.com/kraemer-lab/bayesian\_stats\_epi\_practical/tree/main/assignment) as a guideline. Writing your own code will be rated favourably (instead of simply tweaking the guideline code).

**Task 1:**

The country of Krammeria is suffering from a new epidemic of *probabilis fever*, although some regions have zero cases, other regions present thousands of cases in the last period. This large country has a wide range of temperatures, going from below -2 C˚ to over 20 C˚ (annual averages). These simulated data can be found on the assignment folder: a shapefile (./shapefiles/sim\_country.shp) containing a map with associated temperature and case data, and a separate csv file (sim\_data.csv) containing cases, temperature and regions’ names and coordinates (just in case). Using the simulated data and map answer the following questions.

**Task 1.1:**

Plot a descriptive map of the data (e.g. number of cases per region, temperature, etc.) and define a research question and hypothesis based on the whole dataset (map, cases, temperature).

**Task 1.2:**

Write down a statistical model (equations) based on your question and hypothesis from previous task (Task 1.1). Justy your choices.

**Task 1.3:**

Run prior predictive checks on the previous model (from Task 1.2) and explain and justify your choice of priors. Sample your model and provide convergence checks, interpret convergence statistics, and trace plots (Did the model converge?).

**Task 1.4:**

Assess posterior and posterior predictive distributions, providing relevant plots and summaries. Make inference based on them (Did they answer your question? Did they support the hypothesis?). Plot posterior predictive checks and interpret them.

**Task 2:**

Choose your own count data problem and apply a Bayesian workflow to it (including writing down the formulas for the model). You can either simulate the data or find a small dataset (or subset of a bigger dataset) of no more than 500 observations. Data must be included with the submission.