

# Earth Day Event 2023

## Data Science for Climate Change

4/28/23

### Hackathon Instructions

#### Data

This repository contains three data files in the *NetCDF* format with suffix *.nc*. The files contain data on temperature anomalies for a list of latitudes and longitudes across the globe in a 10 by 10 grid. In all files, the last year (12 observations) at the (6,6) grid is missing. Moreover: \* Ex2 also has missing data in grids (5:6)X(6:7). \* Ex3 has missing data in grids (5:7)X(5:7).

#### The problem

Your goal is to impute/predict/forecast the missing observations at the (6,6) grid. You can use any method and programming language for prediction, but your results should be shown in a CSV (comma-separated values) file.

Alongside your predictions, the code used to create the predictions should be provided.

The winner of the hackathon is the person/team which predictions achieve the minimum squared loss from the true temperature anomalies while using the least amount of information.

We use a point system to determine the winner. Points are awarded in a reverse-order to the top 10 results from solving the exercise using the first dataset *temp\_anomaly\_ex1.nc*. That is, the top solution gets 10 points, second gets 9, and so on.

Once you are confident with your solution to the first exercise, additional points can be achieved by solving the problem using the second or third datasets. Points are awarded in a reverse order for the top 3 solutions for the exercise using the third dataset, and for the top 2 solutions for the exercise using the second dataset.

## The format

You will have 3.5 hours for the task. You can work in groups or by yourself. We target the groups to be at *a maximum of 3 participants*, but bigger groups can be accommodated.

There is some example starter code using Julia in this repository, you should read it first to familiarize yourselves with the file format.

You may ask questions during the hackathon, but **try to google things first** to speed things up.