

# Individual Plan

Course: Data Science for Smart Environments

Student: Hudson Passos (1311182)

MEE (M. Earth and Environment)

## Background and motivation:

I have experience with ArcGIS Pro and QGIS, with some basic knowledge of Python and R. A few years ago, I took a long course on machine learning and deep learning, but I haven't had much practice since then. I chose this course because I think it related to my field of interest which lays in a combination of geospatial analysis with data science. However, I still have several knowledge gaps and I haven't seen yet so many applications in this area. I believe in this course I can resolve both issues pursuing the following goals:

## Goals:

1. Perform geospatial analysis in Python (e.g. geopandas).
2. Combining information from different sources.
3. Manage my scripts and data with github.

## Intertwined/contributes with the project

**(1) geospatial analysis in Pandas:** we drafted a list of spatial data that will be useful for the data analysis related to the causes of air pollution. For instance: population density and land-use (tiff), power plant locations (csv point), administrative boundaries (shapefile), etc.

**(2) combining information:** we tracked many data from different sources that might be somehow connected to produce new information. For example, combining data of air pollution with respiratory diseases can indicate a certain correlation and then a new created information.

**(3) github:** we can share the project scripts with the group and maintain a version control.

## Current level and level to achieve

- **Geospatial analysis in Python:** current level (very low); to achieve (intermediate).
- **Combining information:** current level (low); to achieve (intermediate).
- **Github:** current level (very low); to achieve (intermediate).