

==== EXERCISE 01, FIRST YEAR PROJECT 2022, PROJECT 1

Doing this exercise is doing work for your submission. You do not have to finish all the points listed below within the exercise time of 2 hours! Use the exercise time to start (with help from TAs), and then finish everything within 2 days together with your group. The next exercise will build on that.

SYSTEM SETUP

- Create a Github repository as instructed from the previous lecture
- Make sure you have folium and pandas installed

- Set up a new github project with your group for project 2, you can use github.itu.dk. Call it "fyp2022p01gXX" where XX stands for your group number (01,02,..)
- Set up a minimal github project folder structure. You MUST have:

- README.md
- folder data with subfolders raw, interim, processed, external
- folder notebooks
- folder references
- folder reports with subfolder figures
- folder scripts

- Put the raw data into data/raw.

- While creating your code for the following exploratory data analysis, make several commits and pushes to get a feel for git. Take care to coordinate with your group members, so you don't get into edit conflicts.

EXPLORATORY WEATHER DATA ANALYSIS

- Create a Jupyter notebook in the folder notebooks. In this notebook, make an initial exploratory data analysis, where you:

- load the data
- report the dimensions of the data (number of rows / fields in the weather dataframe)
- make a sanity check: are there missing values anywhere in the weather data?
- filter the weather dataframe so that it only contains rows referring to your country of study
- aggregate the weather data by region and calculate the min, mean, median, and max of all numeric columns. Then plot one of these variables of your choice.
- aggregate the weather data by date and calculate the mean of all numeric columns. Then plot one of these variables of your choice.