Second Assignment

Global Warming

Research question

What is the average rate of temperature change in *your country* since 1960, and how good is a linear regression model?

Learning goals

- Use Python for developing Python modules, i.e. create a separate script for functions that is sourced in the main script (function: import).
- Calibrate a linear model (→ Lecture 2.2 and 7.3), evaluate it (→ Lecture 8.1), and validate it (→ Lecture 8.2).
- Document your code according to a standard style (→ Lecture 9.2).

Data

The data is the same as in assignment 1.

The country assigned to you can be found in the file **student_country_assignment1.csv**.

Temperature anomaly data from Berkeley Earth can be downloaded here. Click on the link to your country, and then on the link to the data table below the top figure. It opens a text file name ***country*-TAVG-Trend.txt**. Right-click on it, and then click on 'save page as'.

Tasks

- Task 1: Read the data into Python, 分组 subset them to start in 1960, calculate annual averages, and 标定、校准 calibrate a linear regression model (e.g., using the linregress function from the scipy.stats library). Print the slope average rate of temperature change in °C / /year乘100 century to the console.
- Task 2: Evaluate your model with at least 3 complementary 辅助标准 criteria (excluding or in addition to those automatically calculated during the model calibration; e.g., percent bias, NRMSE, and the coefficient of determination; → Lecture 8.1). Write this output to a data file.
- Task 3: Cross-validate your model with the same 标准 criteria (e.g., by performing a 5-fold cross-validation, but not leave-one-out; → Lecture 8.2). Write this output to the same data file. Make sure that the data file is easily readable by properly aligning columns.
- Task 4: Plot the time series as well as the linear regression line into one figure, including a well-chosen symbology and conventional annotations such as a legend. Save the figure in a file.
- Task 5:
  (a) Write code that is nice (→ Lecture 3.2, slide 27) and free of errors and warnings (especially warning symbols shown on the left of a Python script in Spyder; → Lecture 5.2).
  (b) Write functions (documented in NumPy or Google style; → Lecture 9.2) into a separate file and call them in the main script.

Deliverables

- The main Python script should be named *country*_assignment2_main.py
- The Python script with function definitions should be named *country*_assignment2_functions.py
- The input data file of your country

When running the script in the same folder as the data input file, no user interaction should be requested. Two additional files should be automatically created and also submitted:

- A tab-separated data file named *country*_statistics.txt
- An image named *country*_time_series.png

→ Collect all the deliverables in a **zip file** to submit it.

Deadline: Friday, **30 November, 14:00** via Brightspace

Assessment

The assignment will be reviewed by 2 of your fellow students. Until **Friday, 30 November, 23:59**, you will **provide feedback to 2 of your peers** and 2 peers will provide feedback to you. Please be fair and honest in your evaluation and provide constructive feedback where you can. The final grading will be done by the lecturer.

The assignment is worth 1/5 of the final grade. The code (syntax and logic), its meaning (semantics), and its output (clear, nice, and relevant) are evaluated.

- Task 1 (calibration): 15%
- Task 2 (evaluation): 25%
- Task 3 (validation): 25%
- Task 4 (plot): 15%
- Task 5 (code): 20%

There will be **no retake** within the same academic year. So, please take this assignment seriously. There is no reason to fail. If you have difficulties with the assignment, seek help on **Google / Stack Overflow** and attend the **Friday workshops**.