Small Python project

1-Project objectives

This project aims at :

- assessing your level of Python
- introduce you to usual libraries for data analysis, if you don't know them yet

  There is a bonus feature of the test which consist in asking you to push your code on Github/Gitlab or similar platform (the choice is yours).

2-Outline of the project

- The data comes from ADEME (to befound on data.gouv.fr) and relates to GHG (Green house Gases) per "communes" in France.
- The data is related to GhG (Green House Gas) in France per "commune". The data, produced by ADEME, are to be found here : https://www.data.gouv.fr/fr/datasets/inventaire-de-gaz-a-effet-de-serre-territorialise/#_.
- The objective is to deliver a small data anlaytics project and use the usual data analytics/data science libraries

2-Instructions and code of conduct

2-1. Instructions

Below are the instructions.

2-1-1. Guidance in Jupyter notebooks

I will provide you with some notebooks to start the project.
It means that the instructions for the project are within that notebook.
You should find enough guidance in the notebooks.
They will be ordered, e.g (illustrative only):
1-Data import.ipynb
2-Data preprocessing.ipynb

Follow the guidance in notebooks to structure your project.

2-1-2. What are my expectations ?

- Be creative and think hard:
  - I am providing minimum guidance BUT I am not spoonfeeding you.
  - Not all steps of the analysis are described, you must think by yourself for intermediate steps that are not explicitly provided !!
  - But don't hesitate to find any way to do more and better !
- Structure of your answer:
  - you MUST submit your project using scripts only (.py files)
  - the only exception is the notebook to present your work !!
  - I am expecting you to have your project follow a proper Python project structure (please justify how you structured your code)

**To make it clear :

- **the results for steps 1-Data import and 2-Data preprocessing MUST be written in .py files.
- **the data analysis part in 3-Data analysis MUST be presented in a notebook.
- Clarification on the notebook to present your project (see 3-Data analysis.ipynb file). I am expecting you to present the following:
  - the structure of your project (you may only replicate the names of the notebooks !).
  - the main sources of inspiration/help (don't lie, you did not produce your answers without help !)
  - the main insights you find useful to share: combine text to comment your analysis with plots.
- Deadline: I request you to share your project at the latest on **Friday 19th April 2024**.
- Submission method: see section below on how you should share your project, so that I can review it.
- Assessment and feedbacks : I will review your code. There are no specific assessment criteria, do your best and JUSTIFY as much as possible your choices.

2-2. Code of conduct

- Help : you'll have access to any help available. I even advise you to use any help that help you answer as well as possible.
- Credits: whenever you use any help, please provide the reference. See below:

- use of ChatGPT or similar : please provide the prompts that you used (i-e the instructions you sent to ChatGPT)
- plagiarism is to be avoided, therefore mention your sources if you have e.g:used StackOverflow or a blog

How to submit your project ?

There are 2 ways:

1- Use of a platform like Github/Gitlab
This is the preferred way to share your code. It will also enable me to hae a look at how you use version control.
Just choose one platform and share the URL of the code repository.
Check that your repository is public so that I can access to the repository.

2-Send me a zipped version of your project via e-mail
It just works, thus I will not penalise you to share via e-mail.

What if you have issues ?

- Search help on the internet (don't forget to mention your sources)
- Use ChatGPT or similar tool in a sensible way (don't forget to mention your sources)
- In any case, be honest, explain your issues and submit what you have produced.