Best practices for your group project structure

Consistent folder structure

To ensure a consistent structure of folders and files in your group project folder, it is recommended to start by setting up a consensus structure all your group members will adhere to. Here is a suggestion:

```
 ~/Documents/t/24/handout_project_structure > tree group_project_name
group_project_name
├── README.md
├── data
│   ├── processed
│   └── raw
├── docs
├── results
├── scripts
└── src
```

- "README.md" is a file containing an overall description of your project
- "data" a folder containing all data of your project .
  - "raw" a subfolder containing the original data you have obtained which you do not want to change.
  - "processed" a subfolder containing the processed datasets you have created.
- "docs" a folder containing any written documentation on what you did, e.g. a draft of your report.
- "results" a folder with the most important visualizations you have created (this could also be stored in "data/processed".
- "scripts" a folder with all your scripts and notebooks displaying each step of your analysis.
- "src" a folder with custom-made modules and functions you are reusing during your analysis.

You have a file that is outdated (or no longer part of your main workflow) but you are not ready to delete it yet? How about moving it into an "archive" subfolder – a cemetery for soon to be deleted files?

Tidy Juyter notebooks

Keep all your Jupyter notebooks tidy. This means:
- Run them from top to bottom only: avoid nested & non-linear cell executions
- Delete or comment out cells you no longer need
- Keep them structured with sections and subsection headers
- Add short and distinct code comments
- Keep them as concise and short as possible

Consistent file naming

Come up with a consistent file naming across all group members avoiding nearly identical file copies with names such as "FINALv2" or "FINALFINALFINAL.doc".
If your analysis mainly consists of consecutively run Jupyter notebooks it might be worth naming them with alphabetic prefixes:
a_metadata_fetch_n_describe.ipynb
b_sequence_fetch_n_preprocess.ipynb
...