# Exercise 2

The following fields can be included in metadata. Discuss each of these fields and colour-mark them according to one of the following categories, based on their relevance for data reusability:

**Essential** (green) - Must-have for FAIR principles (especially reusability).

**Useful** (blue) - Potentially enhances data usability.

**Irrelevant** (yellow) - The data can still be used without this information.

* **Dataset Title** Clear and descriptive title of the dataset or code project.
* **Software Versions** Specific versions of software or tools used in processing or analysis.
* **Ethical Approvals** Information on ethics committee approval (relevant in human/animal studies).
* **Environmental Conditions during Data Collection** Weather, temperature, habitat, or other conditions during data collection.
* **Study Species or Subject** Name of the organisms, species, or subjects studied.
* **Persistent Identifier (DOI)** Permanent link to the dataset/code (e.g., DOI, Zenodo/GitHub link).
* **Repository link**  
  Link to specific repository where the data is located
* **Data Collection Methodology** Description of procedures, protocols, or methods used for data collection.
* **Operating System** OS used (e.g., Windows, Linux) during data collection.
* **File Formats** Formats of the files shared (e.g., CSV, FASTQ, JSON, PDF).
* **Institution** Name of the affiliated organization or university.
* **Principal Investigator (PI) Name** Full name of the lead researcher responsible for the dataset/code.
* **Data Versioning Information** Version history or numbering system used for updates to the data/code.
* **Reproducibility Details** Information provided to allow others to replicate the study or analysis.
* **Principal Investigator address** Street address and office room number of the PI.
* **Date of Data Collection / Code Creation** Time period or specific dates when data was gathered or code written.
* **Quality Control Measures Taken** Checks and filters applied to ensure data/code quality.
* **License Type (e.g., CC-BY, MIT)** Terms and conditions governing reuse, modification, and distribution.
* **Change Log / Version History** Detailed log of updates, fixes, or changes to the dataset/code over time.
* **Units of Measurement** Standardized units used for numeric variables (e.g., meters, grams).
* **Coffee Brand Consumed During Data Analysis** Brand or type of coffee consumed by researchers during analysis.
* **Geographic Location (with coordinates, if applicable)** Specific locations where data was collected, ideally with coordinates.
* **Variable Definitions (Data Dictionary)** Description of each variable or column used in the dataset.
* **Collaborating Researchers** Names of other researchers or groups involved in data/code production.
* **Contact Information** How to reach the dataset/code creator (email, phone).
* **Data Processing Methods** Steps/tools used to clean, process, or transform raw data.
* **Keywords / Tags** Descriptive terms to facilitate searchability and discovery.
* **Preferred Citation Style** Preference for citation format (APA, MLA, etc.).
* **Data Access Restrictions** Any limitations or conditions placed on dataset/code access.
* **Study Design Description** Overview of how the study was structured (sampling design, treatments).
* **Data Provenance** History or origin of the data/code, including how it was generated.
* **Related Publications** Articles or papers associated with the dataset/code.
* **Instrument Calibration Details** Calibration protocols for instruments used in data collection.
* **Funding Sources** Organizations or grants that funded the study or code development.
* **Weather on Data Submission Day** Weather conditions at the time the dataset/code was submitted.
* **Analytical Methods** Techniques applied to analyze or interpret the data.
* **Species Identification Guide Used** Reference used for species or subject identification.
* **Data Collection Protocols** Standard operating procedures or instructions followed.
* **Code Availability (for computational studies)** Link or location where analysis code/scripts are shared.