

# Metadata design

## What is the metadata?

The metadata is a simple `.csv` file that contains the information to (1) find and retrieve the data saved to a remote server and (2) to segment and transform the data to compare experimental conditions. I would recommend recording as many experimental variables as possible, just to be safe. Each row of the `.csv` file is an individual from the experiment. It is mandatory to have at least the following columns: 'machine\_name', 'region\_id', and 'date' with date in a YYYY-MM-DD format. Without these columns the data cannot be retrieved.

| indiv. | machine_name         | region_id              | date              | condition              | sex              | ... | p              |
|--------|----------------------|------------------------|-------------------|------------------------|------------------|-----|----------------|
| 1      | machine_001          | 1                      | 2016-09-01        | A                      | M                | ... | p <sub>1</sub> |
| 2      | machine_001          | 2                      | 2016-09-01        | B                      | M                | ... | p <sub>2</sub> |
| 3      | machine_002          | 1                      | 2016-09-03        | A                      | F                | ... | p <sub>3</sub> |
| ...    | ...                  | ...                    | ...               | ...                    | ...              | ... | ...            |
| n      | machine <sub>n</sub> | region_id <sub>n</sub> | date <sub>n</sub> | condition <sub>n</sub> | sex <sub>n</sub> | ... | p <sub>n</sub> |

Mandatory

Optional

## Top tips.

- Make the metadata exhaustive. Record everything just encase you need information in the future.
- Once you've finished all your replicates, put them in the same `csv` file with a column "replicate" to identify them. It's easier than loading them separately.
- Make your metadata files straight away, as soon as you start your experiment. Time will erode the little details!

Revision #4

Created 22 November 2022 19:05:58 by Giorgio Gilestro

Updated 23 November 2022 09:47:20 by Giorgio Gilestro