# Raw Data Description

**Data Abstract**

| Item | Description |
|---|---|
| DataSet Characteristic | Multivariate |
| Attribute Characteristic | Categorical, Numerical |
| Associated Tasks | Data Smoothing and Classification |
| Number of Subjects | 6 |
| Number of Attributes | 12 (1 to 7, 11, 13, 14 to 16 in Attribute Information Document) |
| Missing Values | Yes |
| Others N/A | |

**Data File Structure Description**

6 Subjects participated in this Experiment.

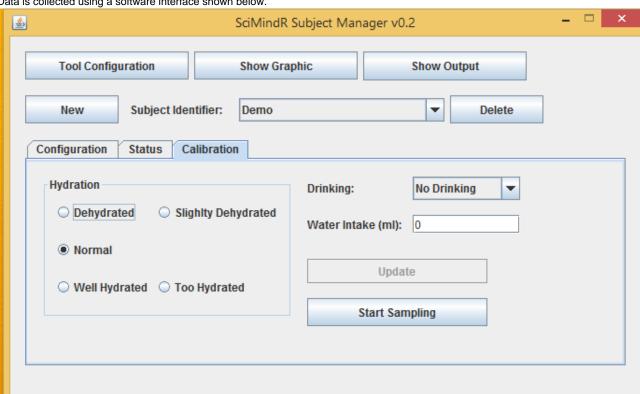Data collected for each subjected are stored in a folder named as "Subject ID".

In the sub-folder of each Subject folder, there are 1 Data folder, in which, all raw data are saved as csv files.

```
Timestamp,Count,Configuration Version,Status Version,Calibration Action
1477425521774,20881,1,1,1
1477425521775,20881,1,1,0
1477425521776,20882,1,1,0
1477425521777,20881,1,1,0
```

(*Each Data File contains Timestamp, Count (RSSI), configuration version, status version, and calibration action.*)

and 3 csv data files corresponding to:

- subjectCalibration.csv
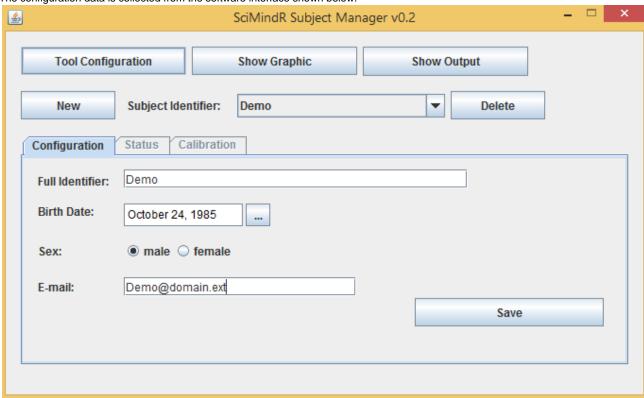  Data is collected using a software interface shown below.

The data looks like:

| Identifier | Sampling Action | Timestamp | Hydration State | Drinking Action | Water Intake | Config Version | Status Version | Exercise Action | Exercise Intensity |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1477425521774 | 2 | 0 | 0 | 1 | 1 | 0 | 0 |
| 2 | 1 | 1477425547874 | 2 | 0 | 0 | 1 | 1 | 0 | 0 |
| 3 | 0 | 1477496944779 | 1 | 0 | 0 | 1 | 2 | 0 | 0 |
| 4 | 1 | 1477497273179 | 1 | 0 | 0 | 1 | 2 | 0 | 0 |

By locating the Timestamp in the raw data file, the hydration status, drinking action, and other activities can be correlated to the corresponding RSSI count.
The Identifier in this table simply corresponds to the Sampling Action.

- subjectConfiguration.csv
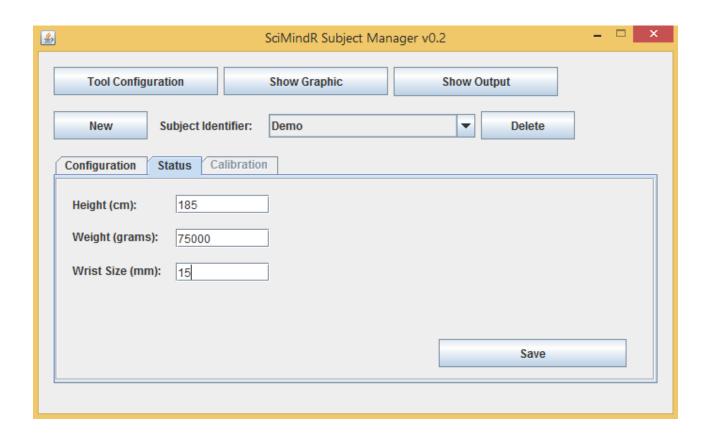  The configuration data is collected from the software interface shown below.



The timestamp information can be used to identify birthdate and gender information, as well as other personal informaiton.

| Version | Timestamp | Identifier | Birth Date | Gender | E-mail |
|---|---|---|---|---|---|
| 1 | 1477425127071 | Sylvain | 1970-10-25 | 0 | Sylvain@domain.ext |
| 2 | 1477582709017 | Sylvain Lizotte | 1970-08-05 | 0 | sylvain.lizotte@biomindr.com |

- subjectStatus.csv
  The following interface is used to collect status information for each subject.

The Timestamp can be used to correlate Height/Weight/Wrist Size information to RSSI readings in the raw data file.

| Version | Timestamp | Height | Weight | Wrist Size |
|---------|-----------|--------|--------|------------|
| 1 | 1477425305225 | 20 | 94100 | 270 |
| 2 | 1477496791482 | 180 | 92400 | 187 |
| 3 | 1477497270000 | 180 | 92400 | 205 |
| 4 | 1477498718790 | 180 | 92400 | 270 |

***Notes regarding real-time Raw Data Observation by participants during test:***

Although experiment is carried out with "minimizing spikes" in mind, the discontinuos spikes in collected data are observed. The likely causes of those spikes could be due to nearby cellphone signals, resonances with external factors, and relocation of device to upper arm etc.

One important data processing step is to "remove" the spikes and keep the "smooth" data to observe the "real" trend in data.