**Data description for paper “How smart do smart meters need to be?”**

**Nataliya Mogles, Ian Walker, Alfonso P. Ramallo-González, JeeHang Lee, Sukumar Natarajan, Julian Padget, Elizabeth Gabe-Thomas, Tom Lovett, Gang Ren, Sylwia Hyniewska, Eamonn O'Neill, Rachid Hourizi, David Coley**

1) File ‘ibert\_study\_QuadStat\_full.csv’ contains data extracted from the ENLITEN project database. This data file was used for statistical analyses of experimental conditions: 2X2 factorial design with factors Value (Environmental Value versus No Value) and Action prompts (Action prompts present versus absent). The file contains each home's daily mean value of four measurements (temperature, CO2 level, gas and electricity consumption) per each experimental condition.

The values in 'type' column define the type of environmental sensor measurements - ‘temperature’ for temperature, ‘co2’ for CO2 level measurements, ‘electricity\_spend’ for electricity consumption, ‘gas\_spend’ for gas consumption. Temperature is expressed in degrees Celcius, CO2 level is expressed in ppm (parts per million), electricity and gas consumption data are expressed in financial costs (pounds £). In order to estimate the electricity consumption in kWh, the standard price of £0.145 for 1 kWh was assumed. Empty cells represent missing values.

2) File ‘common.zip’contains Python scripts used for the data extraction presented in ‘ibert\_study\_QuadStat\_full.csv’ . The files within this compressed file are listed below:

analysis\_quadrant\_full.py (main file)

\_\_init\_\_.py

db\_cfg.py

enliten\_co2.py

enliten\_energy\_stat.py

enliten\_env\_stat.py

enliten\_ibert.py

enliten\_temperature.py

enliten\_visualise.py

ibert\_common.py

iBertPlayingTime.py

mysql\_util.py

navetas.py

pi\_temperature.py

query\_strings.py

3) File ‘finalData.xlsx’ was used for descriptive statistics in Section 5.4 of the paper. The data represent weekly averages of the same four measurements as in file ‘ibert\_study\_QuadStat\_full.csv’. It was extracted from the ENLITEN project database and was cleaned using outliers removal (+/- 3 sigma, and unrealistic zeros, such as temperature=0).

4) File ‘FinalData\_README.docx’ provides a description of the data in the previous file ‘finalData.xlsx’

5) File ‘energy\_literacy\_data\_before.csv’ contains data from a short energy literacy survey conducted prior to digital energy feedback interventions. The survey can be found in Appendix A of the paper.

Description of columns:

|  |  |
| --- | --- |
| Column name | Value description |
| home\_id | Home id as stated in our database |
| q1\_energy | Energy literacy concept, multiple choice, value on [1,4] interval |
| q2\_info | Energy literacy concept, multiple choice, value on [1,5] interval combined with 1 open question |
| q3\_renewable | Energy literacy concept, multiple choice, value on [1,5] interval |
| q4\_uk | Energy literacy concept, multiple choice, value on [1,6] interval |
| q5\_actions | Energy literacy concept, multiple choice, value on [1,5] interval |
| q6\_light | Energy literacy concept, multiple choice, value on [1,5] interval |

6) File ‘survey\_data\_literacy\_satisfaction.csv’ contains data from a combined energy literacy and user experience survey which can be found in Appendix A of the paper. The data was collected after the digital energy feedback interventions. Description of columns:

|  |  |
| --- | --- |
| Column name | Value description |
| home\_id | Home id as stated in our database |
| q1\_energy | Energy literacy concept, multiple choice, value on [1,4] interval |
| q2\_info | Energy literacy concept, multiple choice, value on [1,5] interval combined with 1 open question |
| q3\_renewable | Energy literacy concept, multiple choice, value on [1,5] interval |
| q4\_uk | Energy literacy concept, multiple choice, value on [1,6] interval |
| q5\_actions | Energy literacy concept, multiple choice, value on [1,5] interval |
| q6\_light | Energy literacy concept, multiple choice, value on [1,5] interval |
| q1\_freq | User experience concept, multiple choice, value [1,3] on 3-point Likert scale |
| q2\_dif | User experience concept, multiple choice, value [1,3] on 3-point Likert scale |
| q3\_learn | User experience concept, multiple choice, value [1,3] on 3-point Likert scale |
| q4a\_live | User experience concept, multiple choice, value [1,3] on 3-point Likert scale |
| q4b\_weekly | User experience concept, multiple choice, value [1,3] on 3-point Likert scale |
| q4c\_text | User experience concept, multiple choice, value [1,3] on 3-point Likert scale |
| q5\_useful | User experience concept, multiple choice, value [1,3] on 3-point Likert scale |
| q6\_messages | User experience concept, multiple choice, nominal value: ‘yes, ‘no’, ‘partially’ |
| q7\_actions | User experience concept, multiple choice, value on 2-point Likert scale |
| q8\_more\_actions | User experience concept, multiple choice, nominal value: ‘yes, ‘no’, ‘partially’ |
| q9\_other | User experience concept, multiple choice, nominal value: ‘not aware, ‘aware’, ‘partially aware’ |
| size\_househ | Household size, numerical |
| date\_taken | Date when the survey was taken |
| comments | Comments of a researcher who took the survey |