# Hackathon Data

The data put at disposal for the Hackathon aims to give an overview of the types of information available and can be used as illustration in the apps

The data is split in two parts that can be used in the different competitions.

* “classical” CRM data
* Smart device data

# CRM Data

Consists of 5 datasets that contain information about GDFSuez customers and their interactions. All tables are linked through customer\_id

Local

Contains information about the premise of a customer.

|  |  |
| --- | --- |
| local\_id | ID for the premise |
| customer\_id | ID linking customers |
| Number\_people \_household |  |
| occupation\_type | owner or tenant |
| residence\_type | main home or second home? |
| housing\_type | individual or collective housing |
| property\_type | private or social residence |
| heating\_mode | individual or collective heating |
| heating\_energy | energy used for heating |
| Energy\_hot\_water | energy used to warm the water |

Contact

Contains socio-demographic information about customers, as well as client segmentation and information about IT equipment.

|  |  |
| --- | --- |
| customer\_id | ID linking customers |
| civility | title |
| Name |  |
| gender |  |
| age |  |
| pursuit | job type |
| income | income bracket |
| segment\_client | type of client (Gold,Platinum,Silver) |
| relocation score | a score that indicates whether a client will move or not |
| equipement\_numerique | type of computer (Laptop,Tablet or Smartphone) |
| marque | computer brand |

Balance

Contains the amount that is due by customers.

|  |  |
| --- | --- |
| customer\_id | ID linking customers |
| account\_id | ID linking accounts |
| last\_move | customers last move |
| balance | remaining balance |
| currency | currency of remaining balance (€) |

Communications

Has a record for all interactions between GDFSuez and the customer by phone, web or other.

|  |  |
| --- | --- |
| customer\_id | ID linking customers |
| contact\_id | ID of the contact |
| contact\_category | type of contact (web, phone, …) |
| direction | incoming or outgoing contact |
| start\_date | start of the contact |
| end\_date | end of the contact |

Statement

Has a record for the measurements that were taken at the power and gas meters of the clients

|  |  |
| --- | --- |
| customer\_id | ID linking customers |
| statement\_id |  |
| sector\_activity | Type of Energy (Natural Gas or Electricity) |
| statement\_category | How the measure was obtained (Self identified by the client, Statement by the distributor, estimated by the distributor,.. |
| start\_date\_index | previous measuring date |
| end\_date\_index | current measuring date |
| KWH\_consumption | consumption in KWh |
| unity | unit |
|  |  |

# Smart data

4 datasets are made available with timeseries data coming from smart devices. There is no link between the different datasets.

All data is structured as follows:

* Each row is a different day
* Each column represents a combination of a time in the day and a specific device.

Example below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Device1-0:00 | Device1-0:15 | … | … | Device2-0:00 |
| 1/05/2014 |  |  |  |  |  |
| 2/05/2014 |  |  |  |  |  |
| 3/05/2014 |  |  |  |  |  |
| 4/05/2014 |  |  |  |  |  |
| 5/05/2014 |  |  |  |  |  |
| 6/05/2014 |  |  |  |  |  |
| 7/05/2014 |  |  |  |  |  |
| 8/05/2014 |  |  |  |  |  |

### Smart\_meter\_power

Contains 1 year of data from two metered power consumptions with 15 minute intervals. (Devices M1 and M2)

### Smart\_meter\_gas

Contains 1 year of data from two metered gas consumptions with 60 minute intervals. (Devices M1 and M2)

### Smart\_plugs

Contains 1 month of data for 4 individually measured powerplugs in 15 minute intervals.

Measurements are given for:

* Washing Machine
* TV
* Coffee Machine
* Freeze

### Smart\_Thermostat

A smart thermostat allows to control the functioning of the heating by distance and programmatically.

This dataset contains the desired temperature (setpoint), and the actual temperature (temp) for two devices in 15’ intervals. (Devices M1 and M2)