Sensor Data

Type: Group Assignment (Take home)
Team Size: 4-5 students
Marks: 10
Deadline: March 6, 2018 (23:59 PM)

Introduction

An IOT sensor has been designed to measure 6 environmental parameters such as temperature, humidity, CO_2, VOC (Volatile Organic Compound), Light and noise of a closed room in a commercial buildings. Each IOT sensor posts data to the database. A building may comprise of serval sensors located at certain distance from each other.

Files

Assignment_Data.csv and *Assignment_Data2.csv*

Data Structure

Sample data structure:

date_time	unitid	Temperat	Noise	Light	Co2	VOC	Humidity
1/3/2017 0:00	SS0036	24.20417	52.3591	400.4167	436.25	319.9583	60.4125
1/3/2017 0:01	SS0036	24.2	52.4863	399.6667	433.7917	318.2083	60.43333
1/3/2017 0:02	SS0036	24.18333	52.63881	390.125	434.4583	318.75	60.51667
1/3/2017 0:03	SS0036	24.18333	52.53759	376.2917	432.75	317.3333	60.54583
1/3/2017 0:04	SS0036	24.175	52.84477	399.7083	432.375	317.0833	60.55
1/3/2017 0:05	SS0036	24.175	52.06553	395.2083	432.625	317.25	60.55417
1/3/2017 0:06	SS0036	24.17917	52.10936	391.7083	432.8333	317.4583	60.60833
1/3/2017 0:07	SS0036	24.19167	51.86475	374.1667	433.4167	317.9583	60.6

Meta-Data

1. **date_time** is **timestamp** at which sensor posts the data
2. **unitid** is the IOT sensor ID
3. **Temperature** value is measured in °C
4. **Noise** value is measured in db
5. **Light** value is measured in Lux
6. **Co2** (Carbon di oxide) values measured in ppm
7. **VOC** (Volatile Organic Compound) values measured in ppm
8. **Humidity** value measured in percentage

Possible Data Issues

There are following issues with the data:

1. Sometime the sensors malfunction and reads abnormal values.
2. IOT sensors are posting data to database using Wi-Fi network. Sometimes due to network issues, same data point are posted more than once.
3. Sometime sensors get disconnected with the network and data will not be recorded for that period

Assignment:

1. Come up with a solution (implement it using R) to handle data issues. Give the steps in your document and accompanying code in R script. [**3 Marks**]
2. Analyse the prepared data and submit 10-12 findings in no more than 4-5 sentences or bullet points for each finding. Accompanying code in R script. [**4 Marks**]
3. You must organize and document your code to facilitate understanding. For example, state the need when you create a feature, use appropriate variable names etc. [**1 Mark**]
4. Bonus for creative/additional effort. [**2 Marks**]

Submission:

- Rename the solution script and document (doc or pdf) with your name (such as Prakash_DataPrep.R, Prakash_DataPrep.docx) prior to submission on IVLE).
- Submit only one document and one R script covering all parts of this assignment to avoid confusion.
- Provide names of all the team members in your submission document. Please stick to team size of 4-5 students.
- The deadline is same for both fulltime and part time students.
- Any late submission will be penalized.