Big Data Project with NoSQL DB

Data to be analyzed:

ENERNOC electricity consumption data

Anonymized 5-minute energy usage data for 100 commercial/industrial sites for 2012. Metadata for each site is provided including ID, industry, square footage, lat/lng and timezone. Notice that because latitude/longitude are randomized by a small amount, the exact location of a site on a map may not appear correct (e.g. the point might be over a body of water, etc.)

Can be downloaded from: http://open.enernoc.com/data or https://openenernoc-data.s3.amazonaws.com/anon/index.html

For your information: a time series of electricity consumption is called a “Load Curve” (noted ”LD” in the following).

Objectives of the project:

- **Data integration :** explain how the data has been recovered from the EnerNOC website and how it has been integrated into MongoDB (or any chosen NoSQL DB)
- **Data Modeling:** What data model and representation model should you use in your NoSQL database? Why?
- **Query: define and run the following queries with your chosen NoSQL DB**

    - Try some simple queries (5 simples queries): SELECT queries to explore your data, ORDER BY to sort data, etc. You can use what we have seen in the MongoDB Lab
    - Calculate the sum LD for the 100 sites (timestamp interval: 5 minutes)

- Calculate the average LD by sector of activity (timestamp interval: 5 minutes)
- Calculate the total LD for the 100 sites (timestamp interval: a week)
- Calculate the average LD by sector of activity (timestamp interval: a week)

- **Bonus (to go further: cross reference temperatures and energy consumption):** retrieve the weather data (temperatures in US in 2012) + explore the temperature dependence by sector of activity (correlation between electrical consumption and temperature).

Evaluation Criteria:

- **Report:** clearly explain the objectives, the difficulties, justify the solutions of your approach and present the code developed.

deadLine : March, 14th 2016

To do:

- This priject can be done alone or in a group of 2 people
- Send the list of the members' names by email to: isep.rdi@gmail.com

CC to raja@isep.fr