

CS504 - Principles of Data Management and Mining

Data Analytics Course Project

Team Carbon



Theme

Predictability of Power Utilization based upon Historical Smart Meter Energy Consumption Data in the United Kingdom

Problem Selection

The problem we will try to address in this study is how to exploit the multi-variable demographic dependency of energy utilization and local weather data, to analyze patterns of energy consumption at the household level and make recommendations how to potentially expand it to energy forecasting at the level of a greater metropolitan area that tends to be the energy consumption hog of a geographical region.

Data

From the European capital with the greatest footprint, the city of London that is, electricity consumption data from thousands of homes with their Acorn segmentation based upon affluence, lifestyle and economic-social status were carefully selected and curated in an archive enriched with weather data.

Inspired by a greater European Union effort to decrease carbon emissions and a long-term goal to reverse climate change the data set described above will fuel our analysis.

The data originates directly from Kaggle's data set library located at [1].

Scope

Based on the above Kaggle data set our work will revolve around the topics listed under Inspirations on the same page as they are laid out by the data owner/collector

Therefore, the Scope of our work will be exploring only some of the following paraphrased topics considering the course pressing limits of time.

- Geo-demographical segmentation and descriptive statistics of the consumption daily pattern
- Search for potential of disaggregation from the electricity load curve
- Search for patterns in the consumption results, and potential statistical inferences drawn by exploration of consumption variation in relation to ACORN information

- Forecast the electricity consumption of a household and its perturbation upon addition of an electrical heating system, an EV battery system or explore any other electrical system alteration that would render cost savings.
- Forecast at a global scale (London consumption)

Data Use

Kaggle data sets are public, and they bear no privacy restrictions. Most of the time they are meticulously selected, and they are in an excellent shape, coming with data dictionaries and supporting documentation waiting for analysis.

In this case, although the data set is not perfect, it has been ranked with a usability score of 8.2/10 [silver category] henceforth, we want to believe it is a good fit for all intents and purposes of this short-termed data analysis exercise.

Team Roles

Scrum Master:	Jhony Islam
Product Owner:	Spencer Antonio Marlen-Starr
Developer:	Radha Kanuri
Developer:	Cassidy Laskodi
Developer:	Stavros Kalamatianos

References

1. Jean-Michel D. (2019). Smart meters in London - Smart meter data from London area [Data Set]. Kaggle
<https://www.kaggle.com/jeanmidev/smart-meters-in-london>