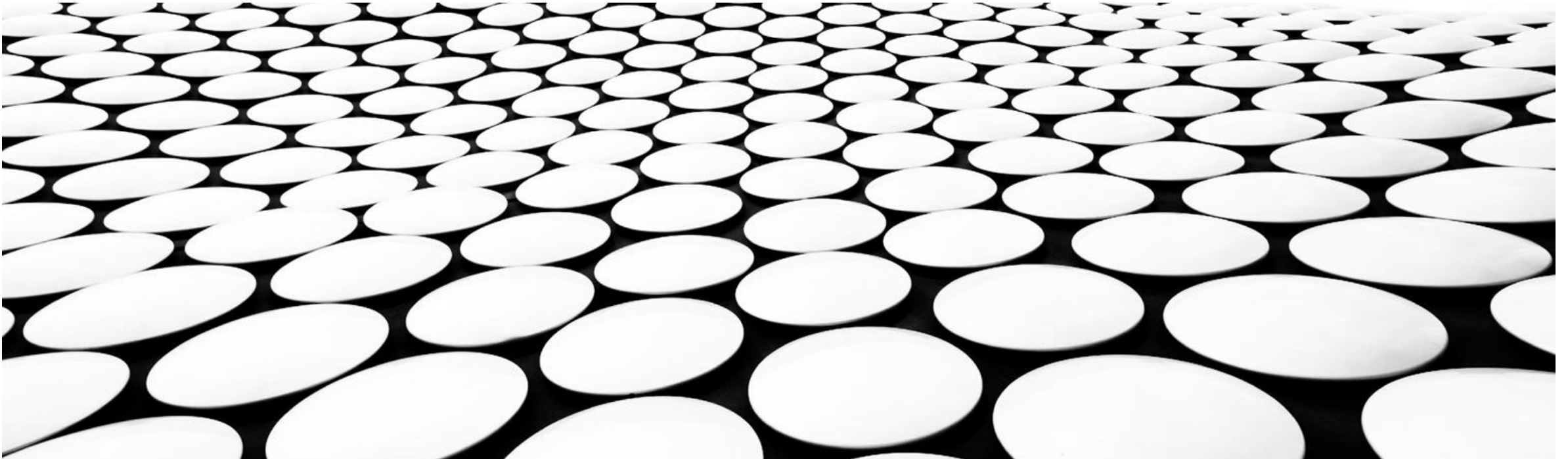# Project Kick-off – Electricity Usage Patterns

Brian Mattis

Agenda

- Background
- Goals
- Data Descriptions
- Data Management
- Known Issues
- Data Statistics
- Recommendations

Background

Photo by Sen Lee on Unsplash

How is electricity tracked?

Utilities track the "real" power draw in each house, where current draw across a voltage differential results in power (W). This is known as the Global Active Power.

Submeters are used to locally track electricity usage of certain subcircuits – such as apartment units or subsegments of a house.

What are submeters?

Submeters are generally used in multifamily buildings to enable the building owner to attribute the electricity consumed by each unit. Submeters are owned/installed by the building owner, rather than the utility company, enabling the building to purchase electricity in bulk at significantly lower rates.

In this case, we have 3 submeters tied to a specific house.

Objectives / Goals

- Objective: IOT Analytics is tasked with analyzing electricity patterns for a specific residence, with implications in a legal dispute.

- Goal: Model the energy usage patterns and investigate how the electricity usage in the summer of 2008 compares with the model. Ultimately, the customer seeks insight on if the electricity usage in the summer of 2008 could indicate the residence was unoccupied for a period of time.

Data Descriptions

- We'll focus our analysis on the 3 submeters attached to the residence and the "global active power" (total power)

- Each submeter records energy usage (watt-hour) in 1 minute increments, 24 hours a day, 365 days a year.

- 47 months of power usage data is available

Residence circuit allocations
- Submeter 1: Kitchen
  - Dishwasher
  - Oven
  - Microwave
- Submeter 2: Laundry Room
  - Washing Machine
  - Dryer
  - Refrigerator
  - Light
- Submeter 3: Heating/cooling
  - Water heater
  - Air Conditioner

Data Management

- Electricity data is retrieved directly from a secure AWS server.
- Data is stored locally on a hard drive for analysis
- Security: Local computer equipped with latest
  - Antivirus software
  - Operating system updates
  - Software updates
- Local computer operates on a secure VPN network


Photo by FLY:D on Unsplash

Known Issues with Data

- The data is incomplete for both 2006 and 2010, missing months of data
- While every date from 2007-2009 contain data, there are 180 missing data points
  - Unknown root cause for missing data points
- Data requires conversion into proper "date/time" formats

Descriptive Statistics

- Average electricity usage is **>4X higher** in Submeter #3 versus other submeters
  - Responsible for the most power usage
- Submeter #3 has a unique **bimodal distribution**
- The most common state for SubMeters #1, #2 is an **"off" state** with zero usage

	SubMeter #1	SubMeter #2	SubMeter #3
Min	0	0	0
1st Quartile	0	0	0
Median	0	0	1
Mean	1.159	1.343	6.216
3rd Quartile	0	1	17
Maximum	82	78	31
Standard Deviation	6.288	5.972	8.341
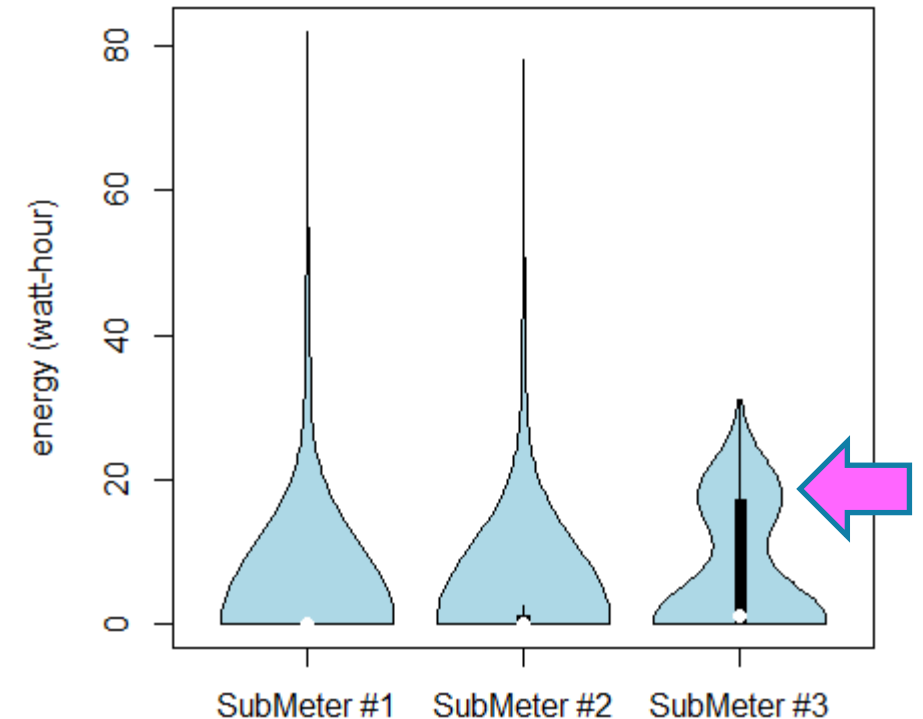
Recommendations

- Secondary data:
  - Include outside temperature data for the client's address to better understand trends in A/C usage (Sub-meter #3)
  - Include the client's work schedule
- Grouping:
  - It would be beneficial if the refrigerator were grouped with sub-meter#1, like the other kitchen appliances
  - It would be beneficial if the electric water heater and the air-conditioner were on separate sub-meters. A/C can be passively set using a thermostat, but hot water usage is more active (showers, dishes, etc)
- We recommend continuing to dive deeper into this data to find meaningful information to the case

Thank you, and we look forward to providing valuable analysis for this case!