# MACHINE LEARNING WITH DOMESTIC ENERGY USE DATA

SAM STERN

# MOTIVATION

- Government plans on installing SMARTMETERS throughout the UK by 2020
- Readings taken by SMARTMETERS stored on a database
- What are the security implication?
  - Taking on the role on an attacker who has hacked into the database and stolen electricity readings.
  - What can you infer from the data?

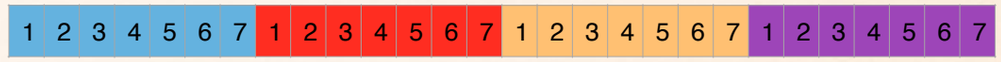
# HES DATASET

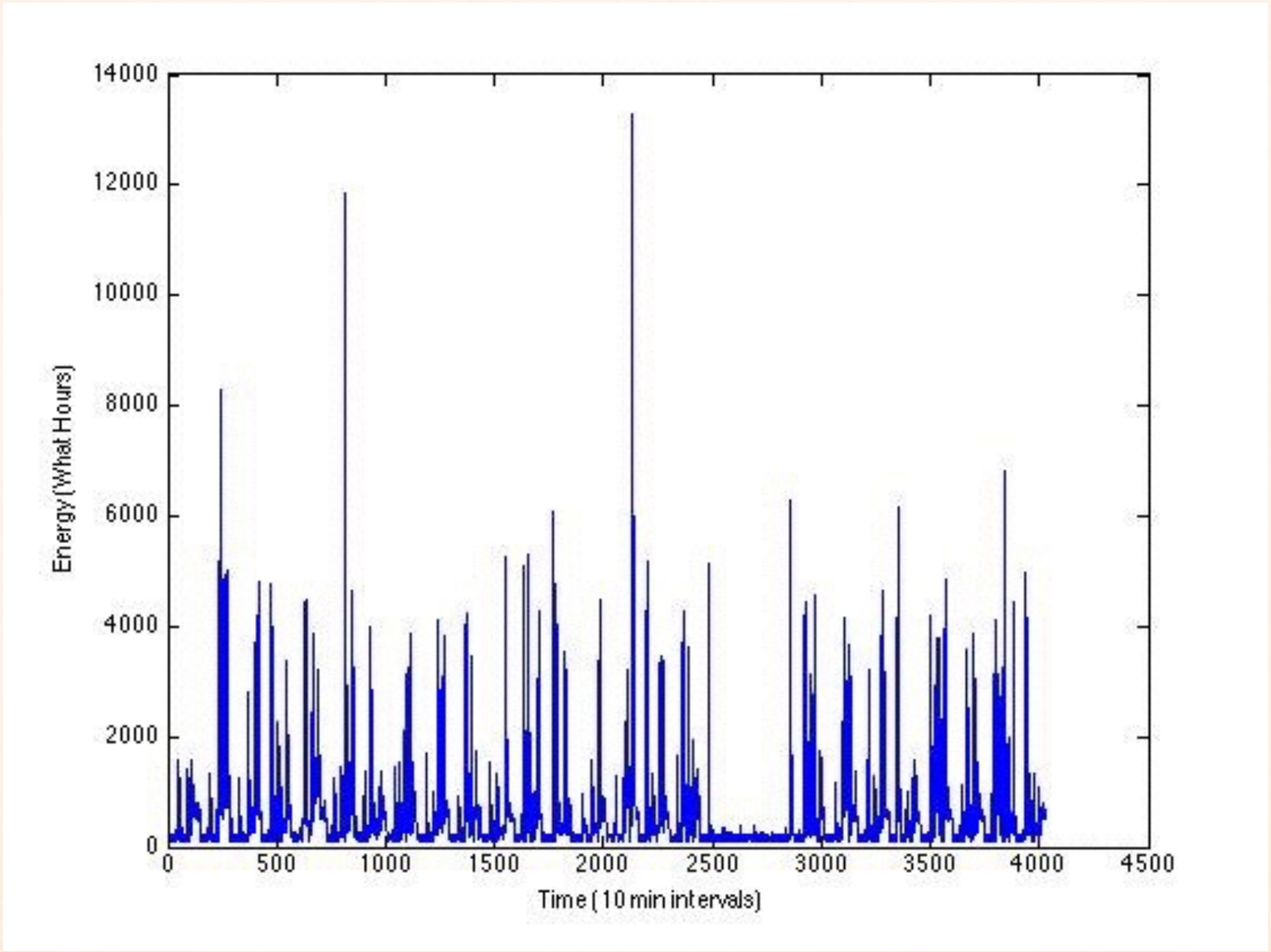
- 250 households
- 26 for 1 year
- 224 for (roughly) 1 month
- Energy measured in either 2 or 10 minute intervals

# PREPROSSESSING

- 1 month
  - easy to implement
  - loost too much data
- 4 weeks
  - What about weekly patterns (energy use different on weekends)
  - for 1 month
    - chop top off
    - it remaining data <28 days
    - recycle/re-use data
  - For 1 year
    - chop top off
    - cut into 28 day intervals




14000
12000
10000
8000
6000
4000
2000
0
Energy (What Hours)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
Time ( 10 min intervals)

# STILL TO DO (BEFORE NEXT SEMESTER)

- Preprossess the rest of the data
- Feature Extraction
  - how fine grained does the data have to be to see distinctions between classes?
  - Use mean, variance, sum, FT or a combination as feature vectors?