



IOT ANALYTICS



Household Electric Power Consumption

Exploratory Data Analysis (EDA) Review

John T. Leonard

Household Power Consumption | EDA : Agenda

- Background
- Objectives
- Data Management
- Statistical Overview
- Time Series Trends
- Recommendations & Next Steps



Household Power Consumption | EDA : Background

- Problem Statement:
 - IOT Analytics closing deal for new sub-metering devices in Smart Homes.
 - Smart Home owners need to understand and control their power usage from these devices
 - Pre-existing power usage data can be used to develop preliminary analytical tools & visualizations for customers.
- Data Set Analyzed:
 - “Individual Household Electric Power Consumption Data Set”
 - Repository: UC Irvine Machine Learning Repository
 - Source: G. Hebrail & A. Berard, EDF R&D, Clamart, France



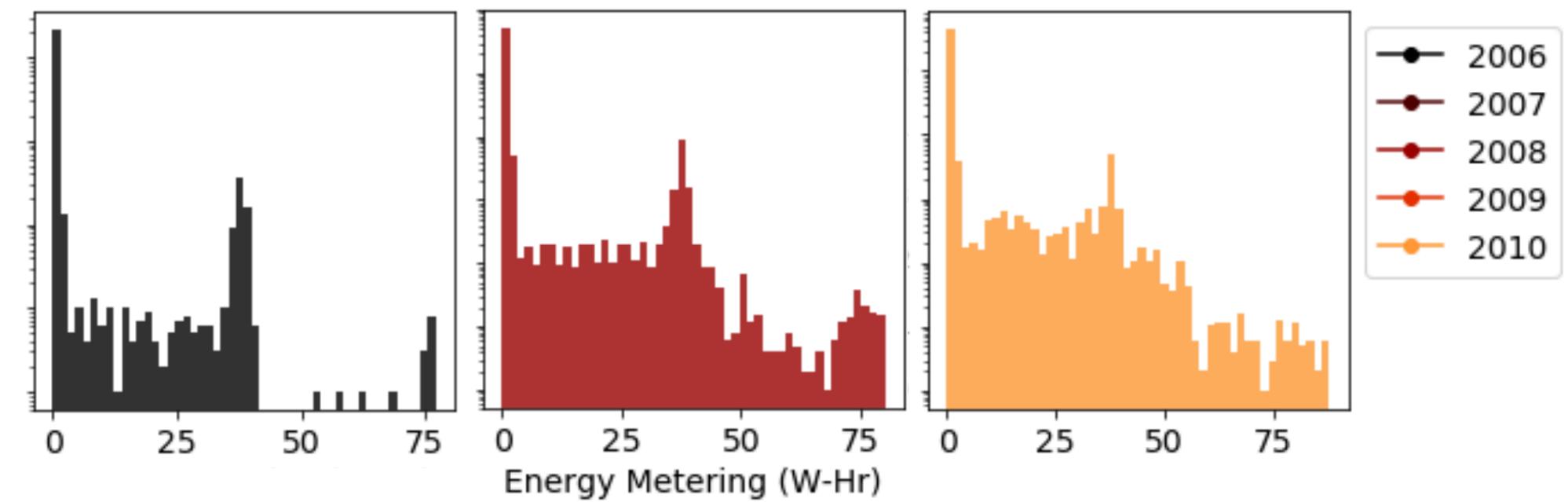
Household Power Consumption | EDA : Objectives

- Objectives

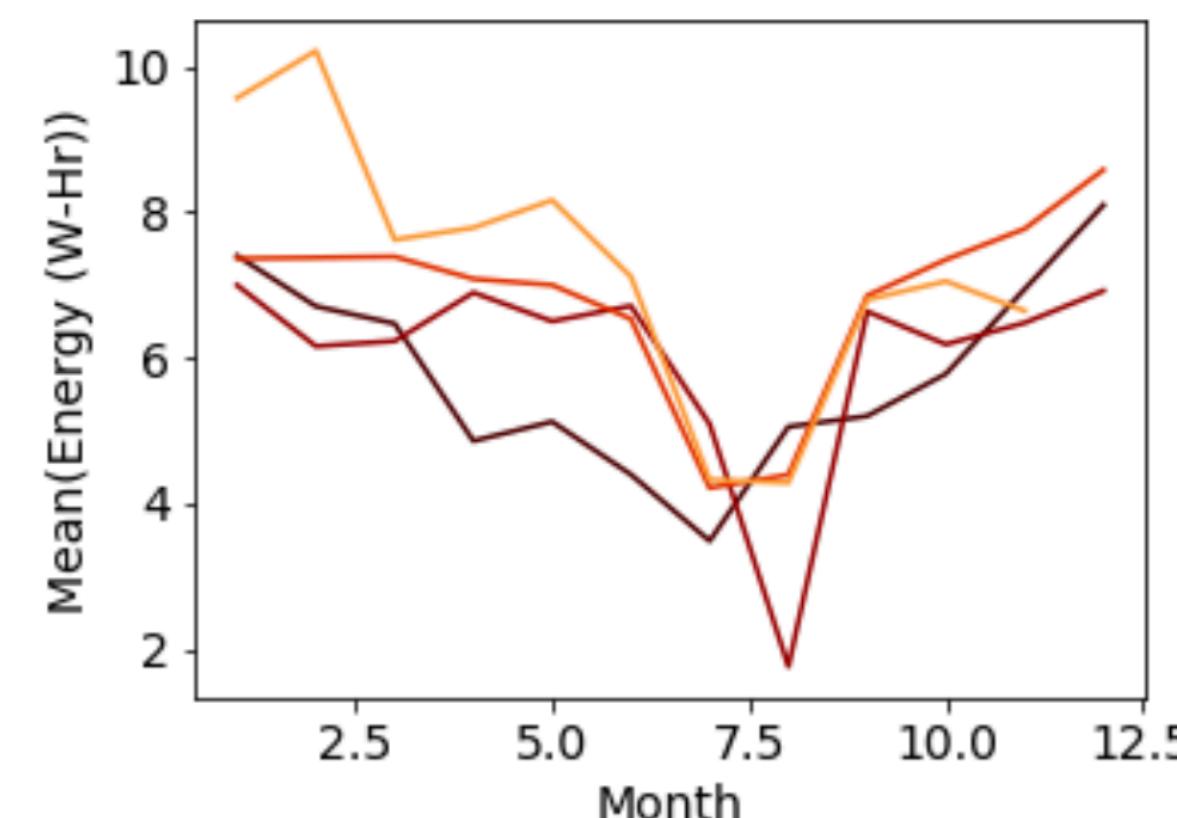
1. Demonstrate SQL Data Compiler Pipeline



2. Statistical Overview of Data



3. Time Series Trends & Visualizations



4. Opportunities for improvement

Household Power Consumption | EDA : Data Management

- SQL: Structured Query Language
 - Data stored on Amazon Web Services (AWS) cloud platform
 - Allows large number of user data queries for large or small volumes of data
 - Queries based on industry standard scripting nomenclature (SQL)
- SQL data compiler pipeline: Downloads & combines power consumption data from 2006 - 2010
 - measured minute by minute
 - 3 sub-meters:
 - #1: Kitchen (Dishwasher, oven, Microwave)
 - #2: Laundry Room (washing-machine, dryer, refrigerator, light)
 - #3: Heating + Cooling (Electric wafer heater, AC unit)



SQL Data Compiler Pipeline

Connect to Database (db)

Examine names of tables

Examine headers for table columns

Download tables+columns of interest

Merge Tables into single database

2006 - 2010 Power Consumption

Sub-Meter 1



Sub-Meter 2

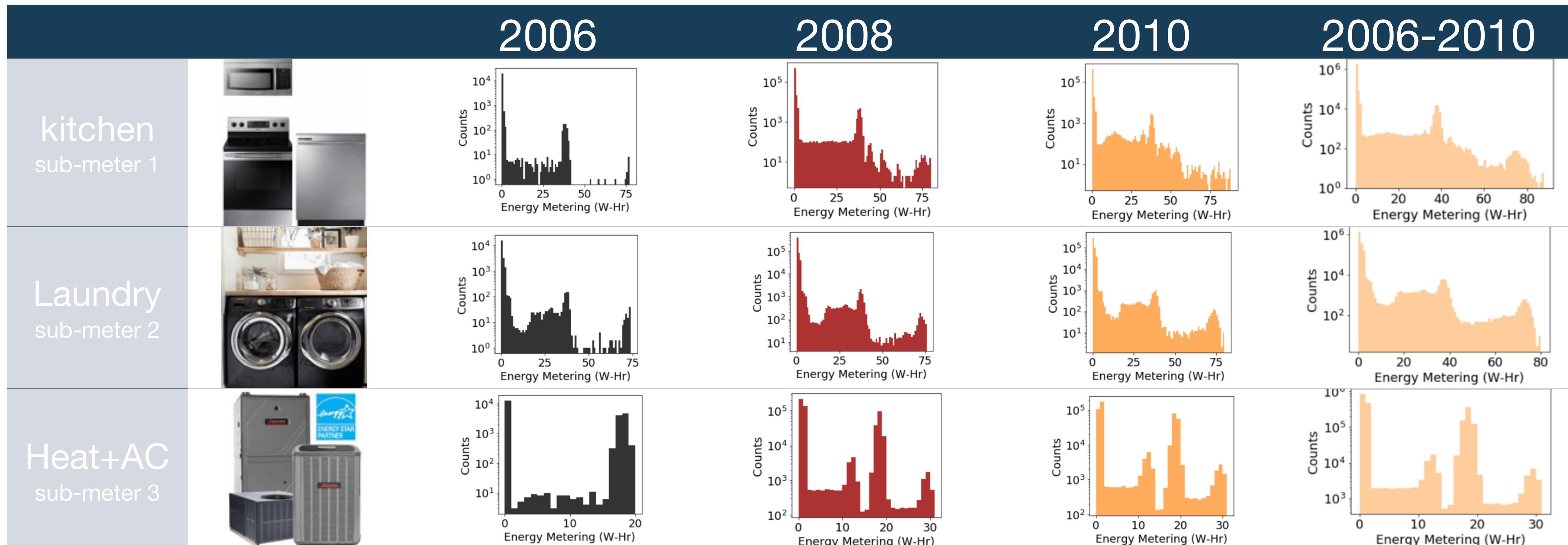


Sub-Meter 3



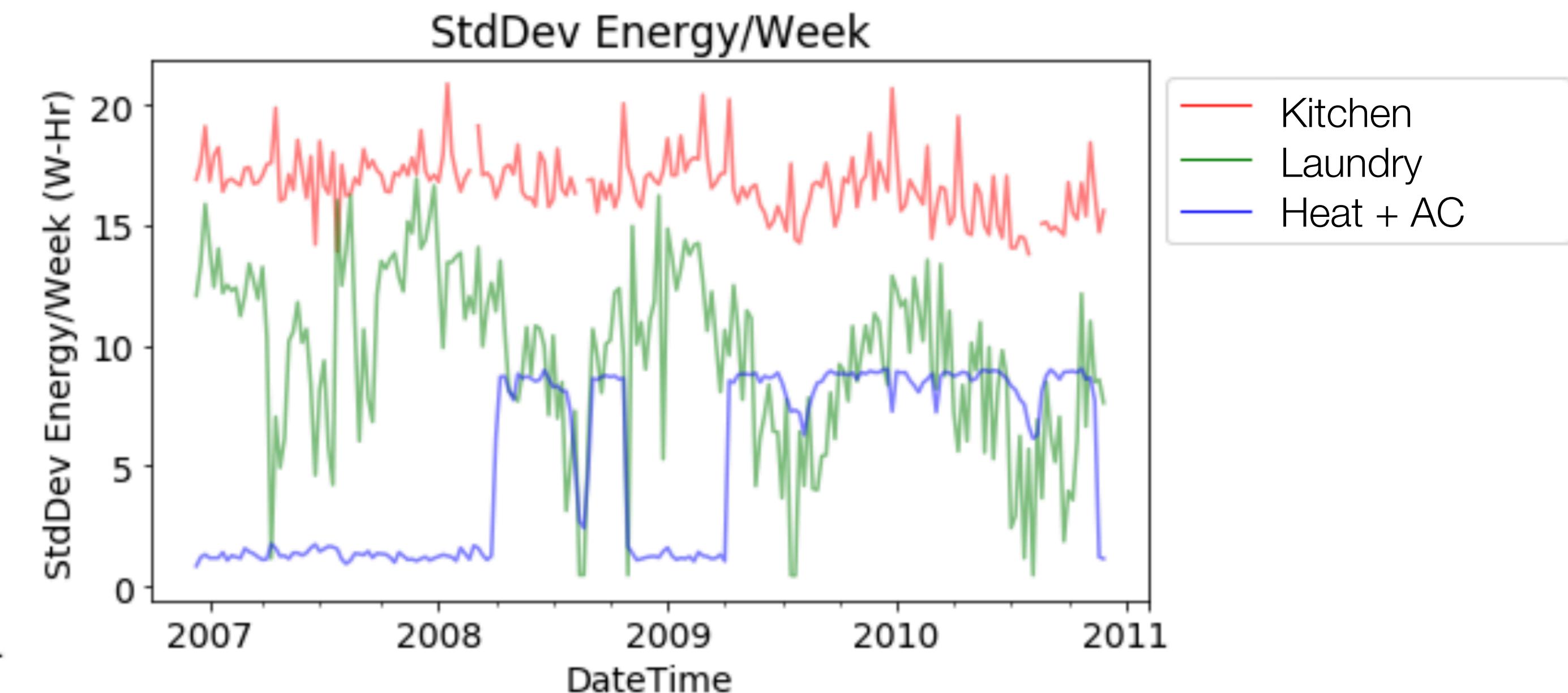
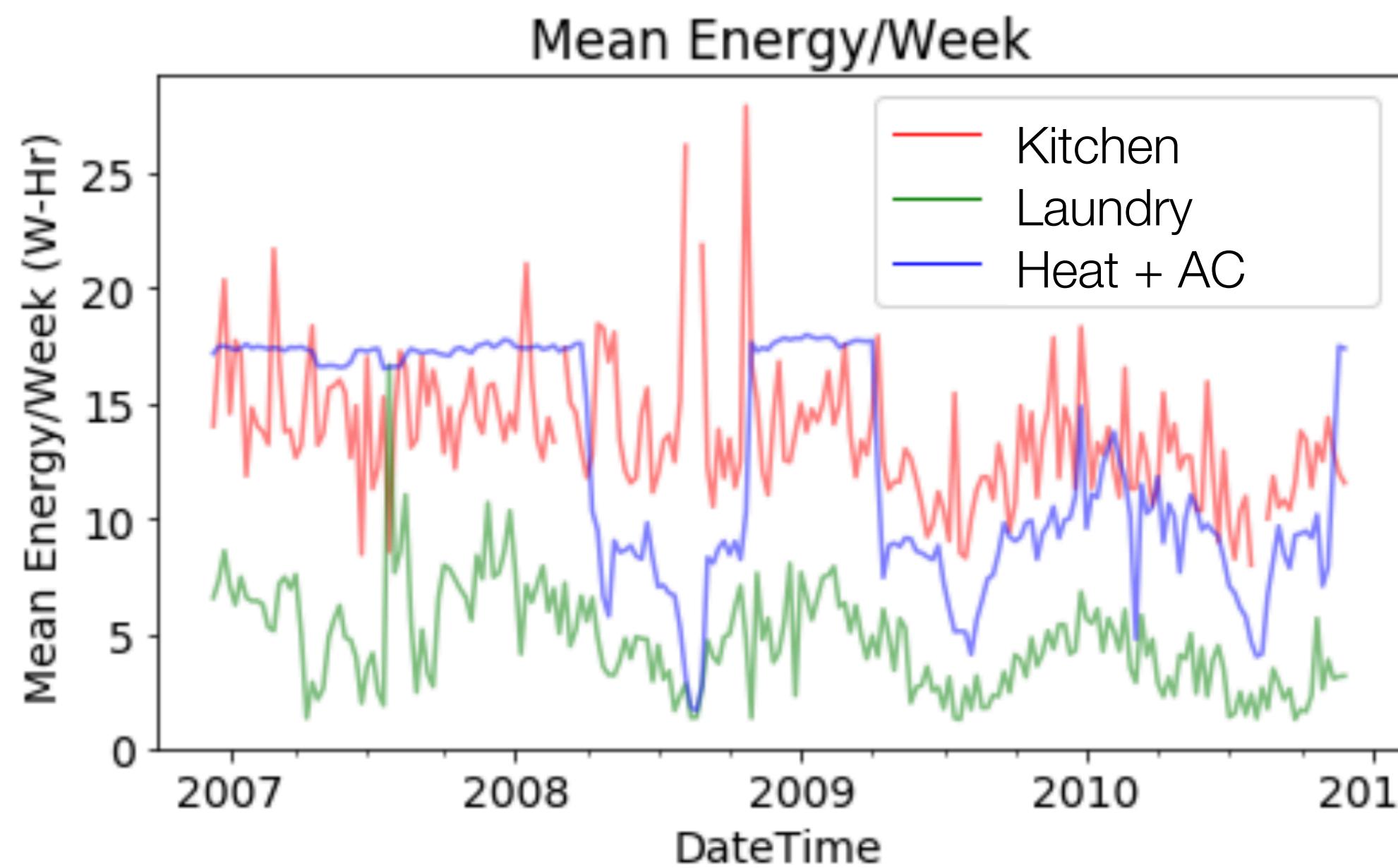
Household Power Consumption | EDA : Statistical Overview

- Histograms for 2006, 2008, 2010 + 2006-2010 (cumulative) power consumption shown
- 3-4 “peaks” in power usage per sub-meter
 - Focusing on cumulative (2006-2010):
 - Kitchen & Laundry: Peaks @ ~0, 40, 78 W-Hr
 - Heat + AC: Peaks @ ~0, 13,19, & 29 W-Hr



Household Power Consumption | EDA : Time Series Trends

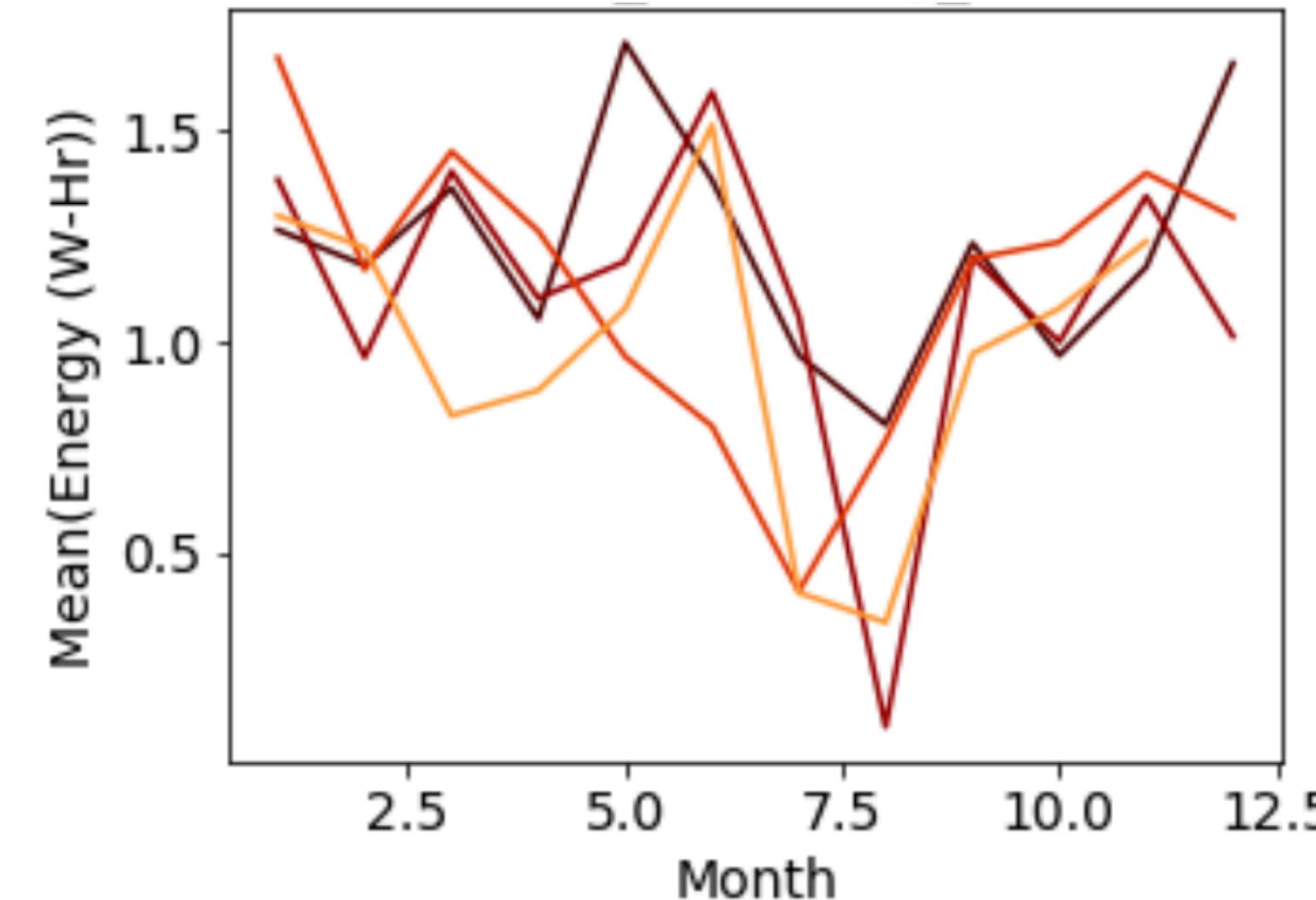
- Drop 0 W-Hr power consumption data
- Plots: Mean & Standard Deviation (StdDev) of power consumption per week vs. time for 2006-2010
- Power consumption of kitchen & laundry reducing over time (more energy efficiency?)
- mean Heat + AC power varies widely year to year.
 - Higher mean Heat+AC power implies lower variance in usage



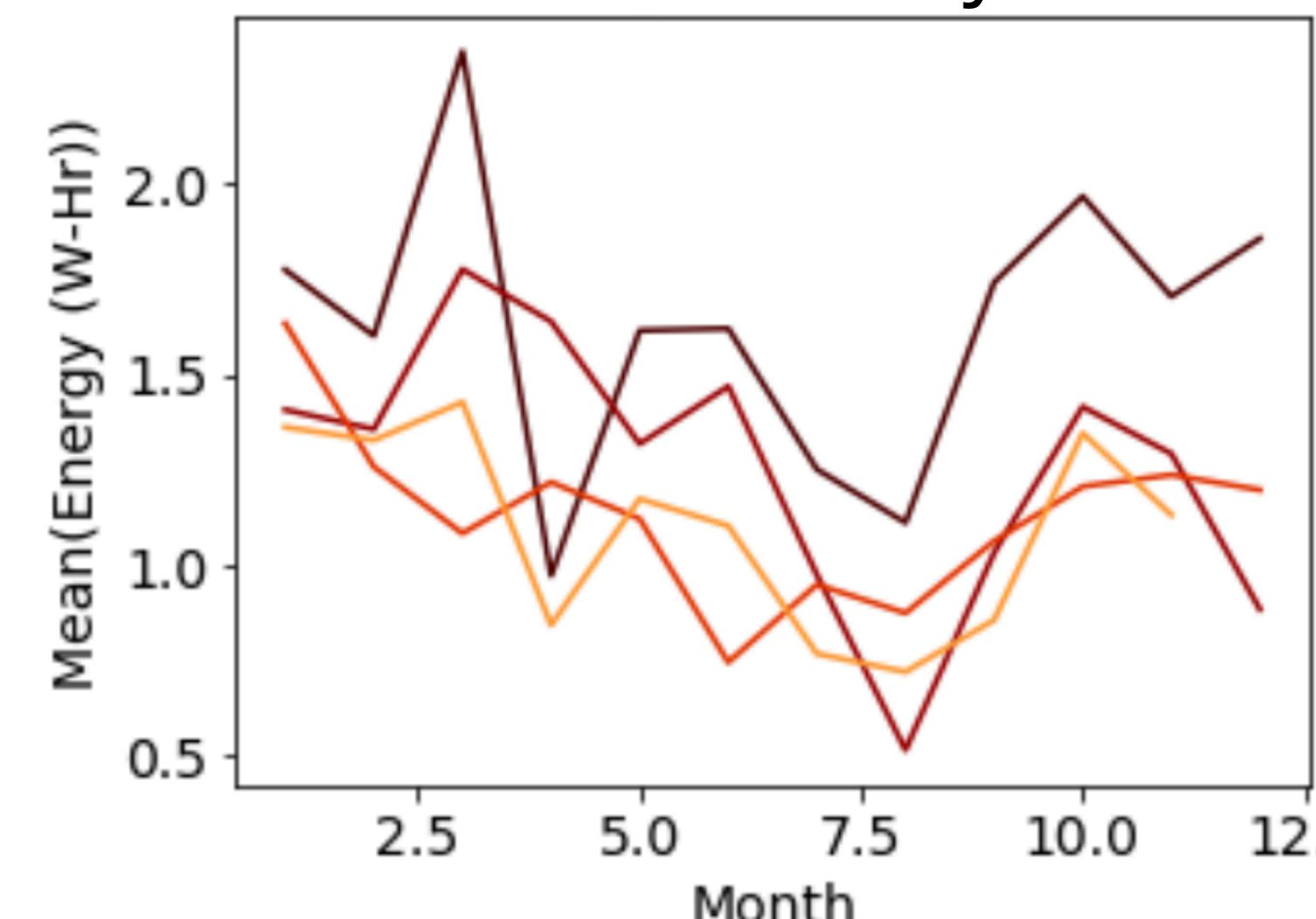
Household Power Consumption | EDA : Time Series Trends

- Seasonality Plots: mean energy per month year by year
- Power consumption drops to lowest in Aug. & highest in Dec./Jan.
- Kitchen & Heat+AC exhibit strong seasonality

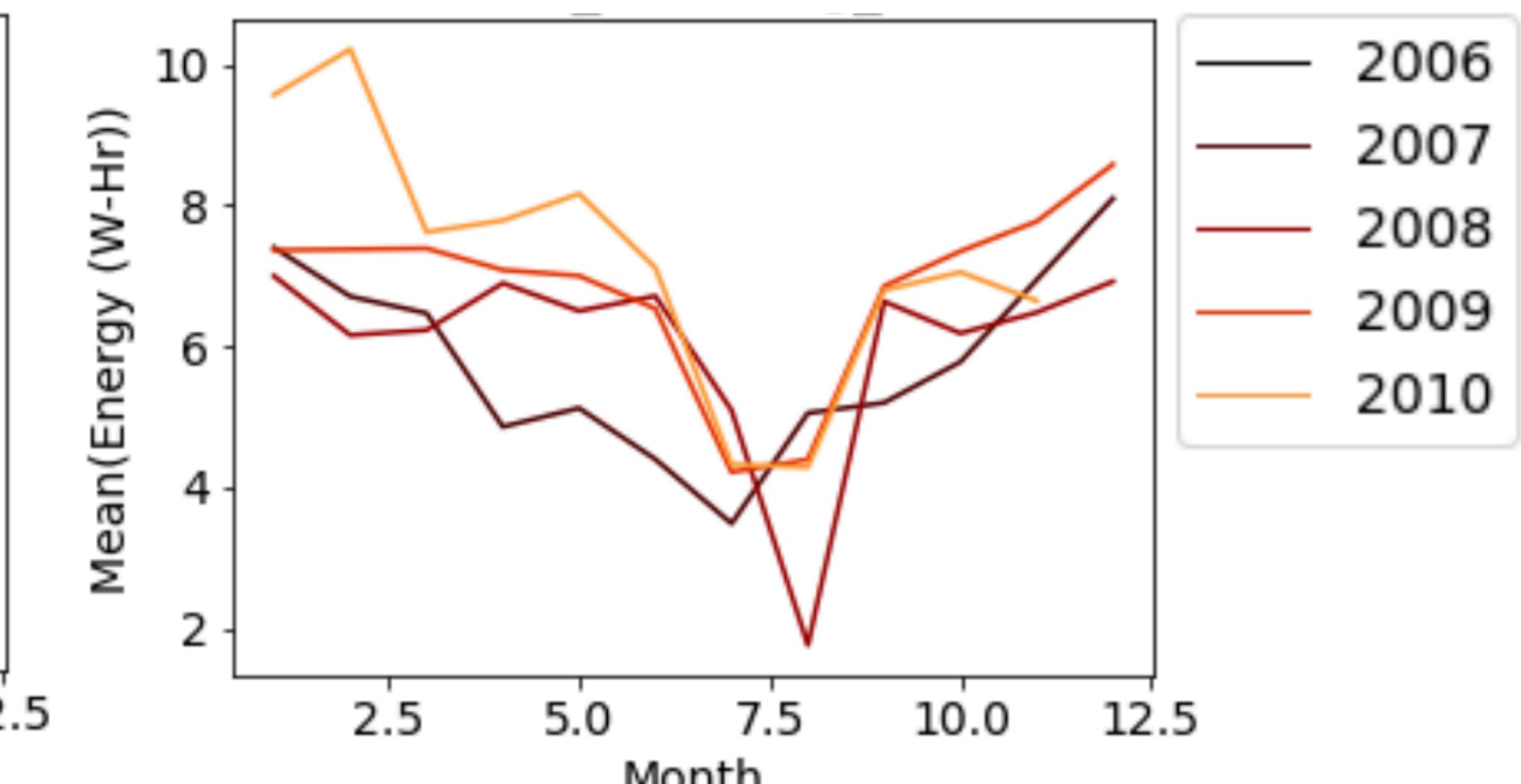
Kitchen



Laundry

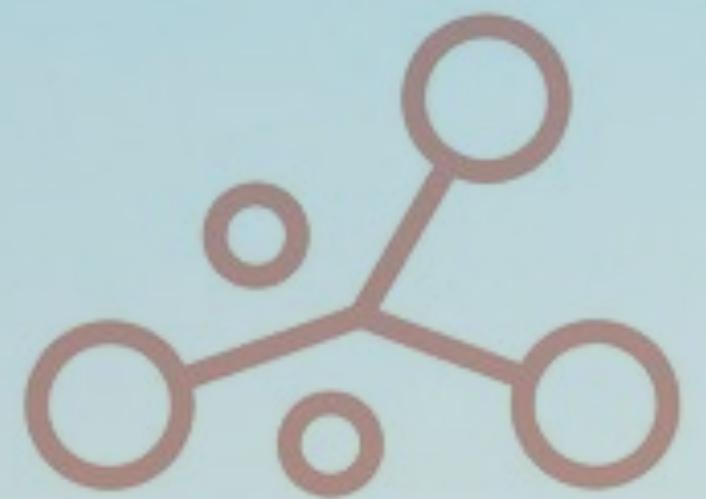


Heat + AC



Household Power Consumption | EDA : Recommendations

- Recommendations:
 - 3 key visualizations
 1. Histogram of power usage per year and/or total
 - identifies typical power usage pts (“peaks”)
 2. Mean+variance of power usage over time
 - For individual customer, could overlay the customers trend with global trends
 3. Mean power per month per year
 - Highlights seasonal trends. Could overlay individual customer trend with global trend
 - Opportunities for improving data
 - Add customer ID as column. Could enable excluding outlier customers
 - Add geographic information (latitude, longitude, etc.): analyze power consumption correlation with geographic location
 - Add weather information: analyze power consumption correlation with weather conditions
- Next Steps:
 - Develop model to predict power consumption over time.



IOT ANALYTICS

