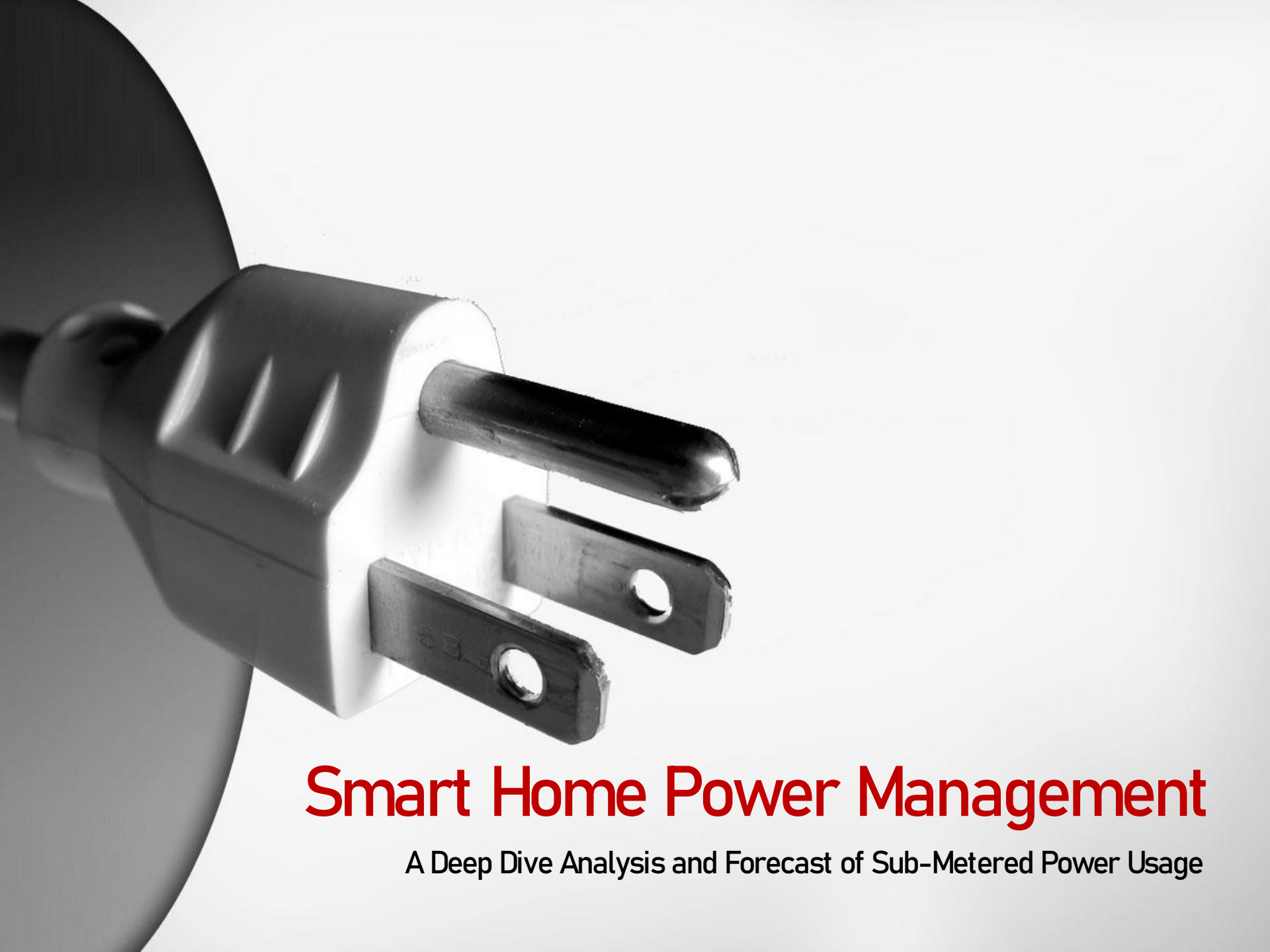

Smart Home Power Management

A Deep Dive Analysis and Forecast of Sub-Metered Power Usage

Project Mission

- Perform a deep dive analysis of smart home sub-metering devices to determine if patterns of energy usage can be identified.
- Demonstrate an ability to forecast future usage based on historical energy consumption.
- Make business recommendations for the client supported by the data.

Data Visualization

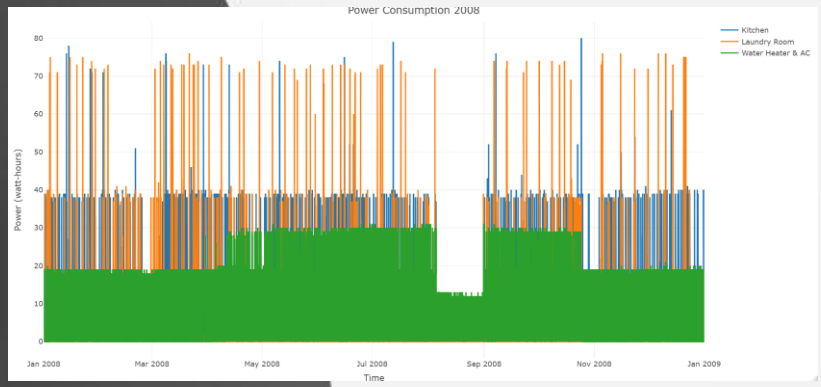

Figure 1: Full Year Power Consumption (one read per minute) - 2008

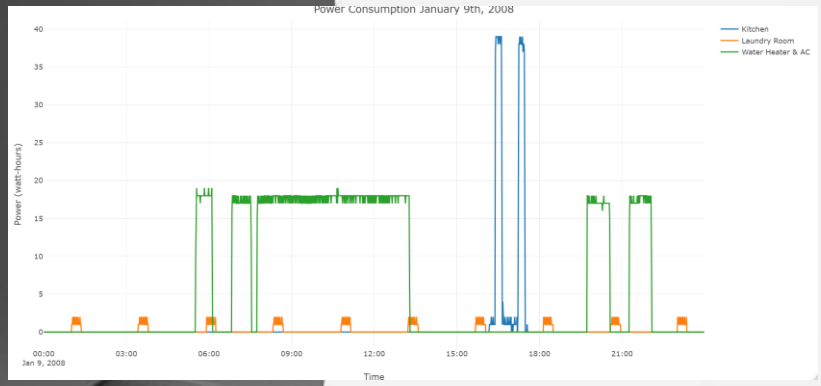

Figure 2: Single Day Power Consumption (one read per minute) – 01/09/2008

Figure 3: Single Day Power Consumption (one read per 10 minutes) – 01/09/2008

- The data consists of individual meter readings taken on 3 sub-meters every minute over the course of 3 years.
- Figure 1 plots every data point over the course of a single year (2008). The graph contains a lot of information but is practically illegible.
- We can break the data down into single days to visualize power usage over the course of an entire day, but that still results in some fuzzy points on the graph (Figure 2).
- If we take a more granular approach, we can plot a reading every 10th minute. We start to get a cleaner graph at the expense of some data loss (Figure 3).

Data Visualization

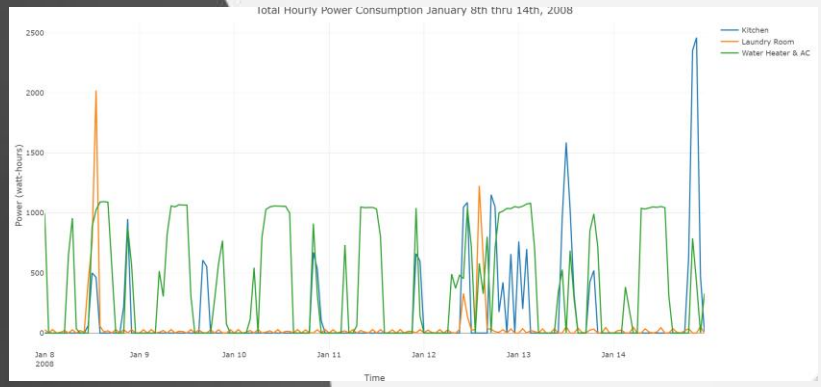

Figure 4: Single Week Power Consumption (total usage per hour) – 01/08-01/14/2008

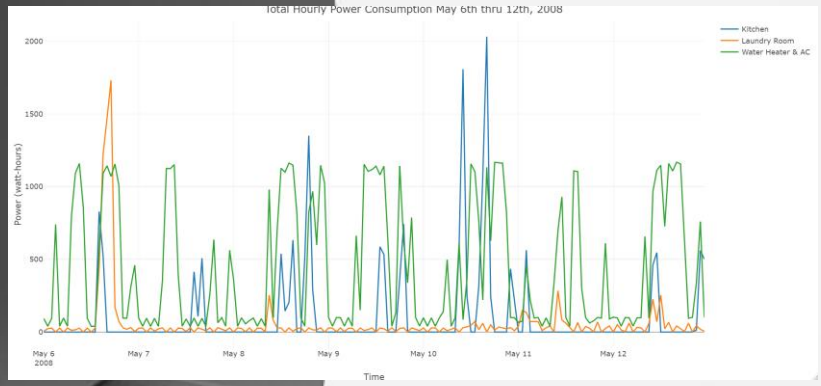

Figure 5: Single Week Power Consumption (total usage per hour) – 05/06-05/12/2008

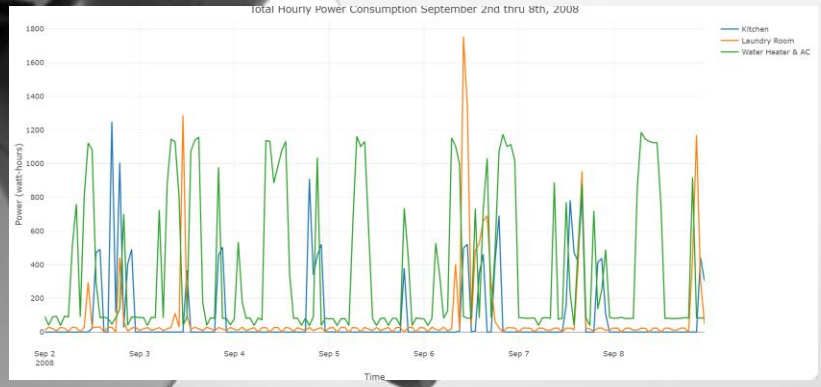

Figure 6: Single Week Power Consumption (total usage per hour) – 09/02-09/08/2008

- We can aggregate the data to see total energy consumption over a given span of time. Figures 4-6 display the total energy usage on each sub-meter every hour over the span of a week (sampled across one year).

- Doing so gives us a clearer distinction of the patterns of power usage by each meter. Sub-meter 1 displays periodic short bursts of energy usage; Sub-meter 2 displays a consistent low level of usage with occasional spikes; Sub-meter 3 displays sustained moderate energy usage at somewhat regular intervals.

- This approach to visualization still only provides us with small snapshots of consumption patterns. What if we want a panoramic view?

Data Visualization

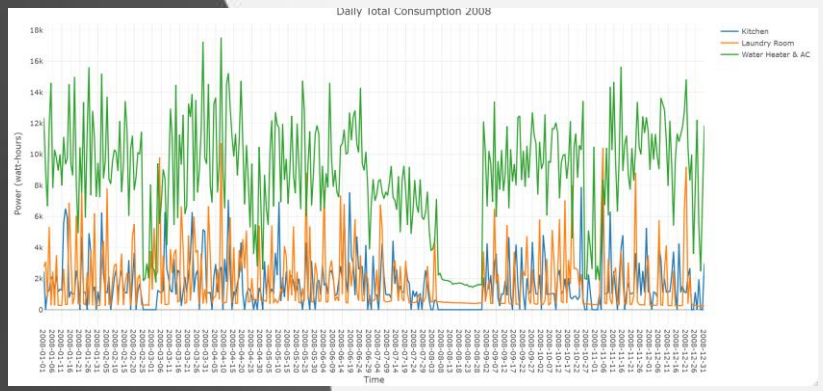

Figure 7: Full Year Power Consumption (total usage per day) – 2008

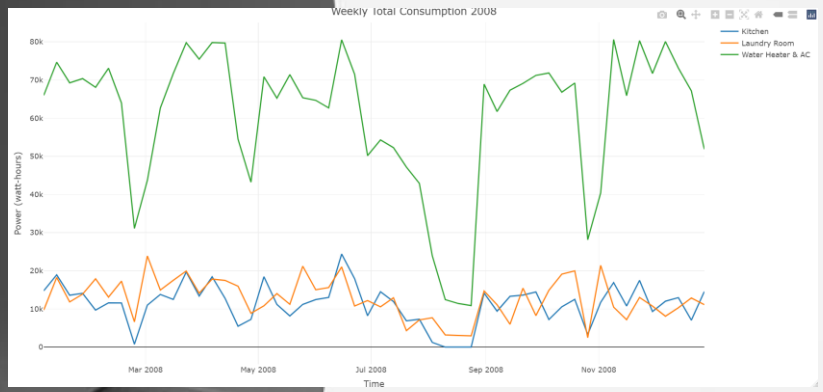

Figure 8: Full Year Power Consumption (total usage per week) – 2008

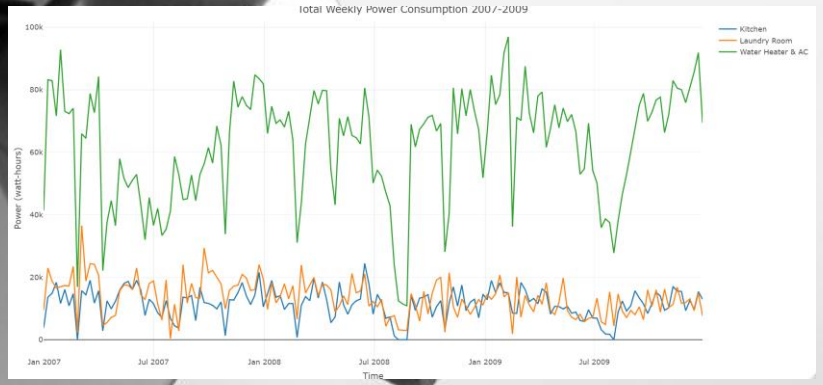

Figure 9: Three-Year Power Consumption (total usage per week) – 2007-2009

- Plotting daily total consumption over the span of one year (Figure 7) begins to give us a clearer picture of sub-meter utilization and a sense of seasonal fluctuation. We can see a substantial drop off in consumption during the summer months, particularly on Sub-meter 3. August sees a near complete halt in consumption before abruptly picking back up to regular levels in September. This may indicate a period of vacation where the homeowner was away.
- Daily data points still create a bit of a cluttered plot, so reducing reads to weekly totals gives us a better summary of usage throughout the year (Figure 8).
- We can use weekly reads to plot a three year history of consumption, giving us a better understanding of seasonal patterns over time (Figure 9).

Time Series Analysis

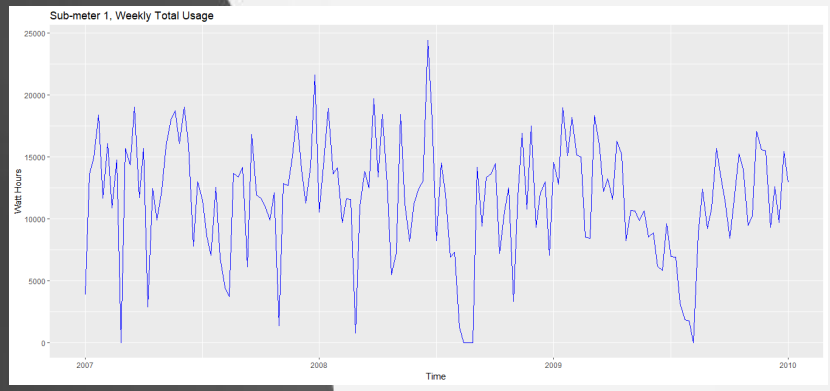

Figure 10: Time Series Plot of Weekly Total Consumption on Sub-Meter 1 – 2007-2009

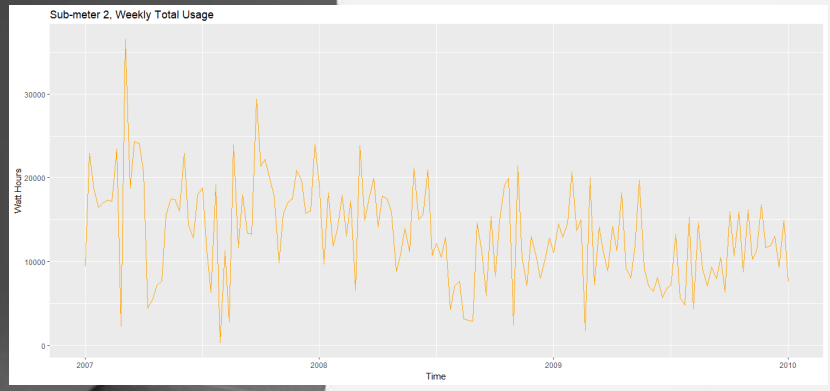

Figure 11: Time Series Plot of Weekly Total Consumption on Sub-Meter 2 – 2007-2009

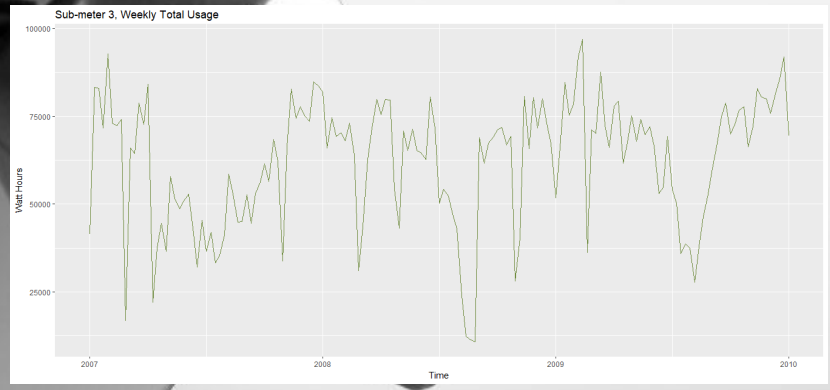

Figure 12: Time Series Plot of Weekly Total Consumption on Sub-Meter 3 – 2007-2009

- Once we determine an appropriate level of aggregation (total consumption) and granularity (weekly reads), we prepare the data for forecasting.
- Upon initial assessment, it is apparent that each sub-meter displays different trends over the studied three year period.
- Sub-meter 1 (kitchen, Figure 10) appears to maintain a mostly level trajectory, albeit with wide fluctuations in usage.
- Sub-meter 2 (laundry room, Figure 11) displays a clear downward trend.
- Sub-meter 3 (AC/water heater, Figure 12) displays clear seasonality in usage and appears to be on an upward trend.
- NOTE: Figures are plotted to illustrate the fluctuation in usage relative to the individual sub-meter but give a false sense of scale relative to one another.

Forecasting – Linear Regression

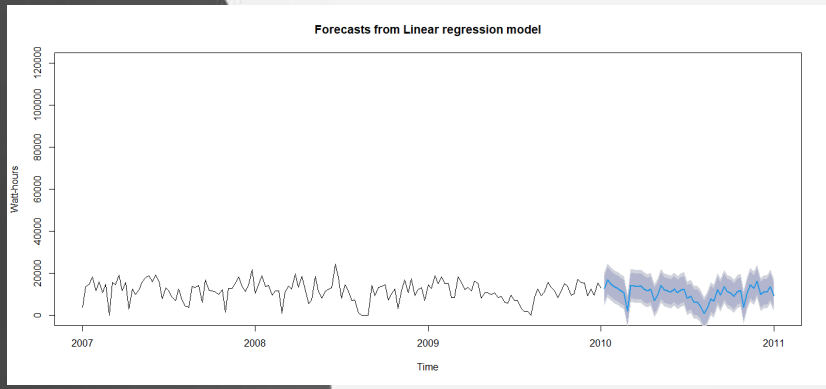

Figure 13: Linear Regression Forecast of Sub-Meter 1 Usage for 2010

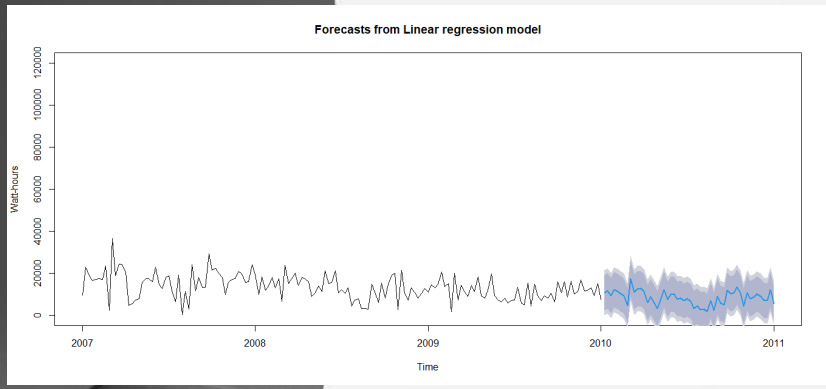

Figure 14: Linear Regression Forecast of Sub-Meter 2 Usage for 2010

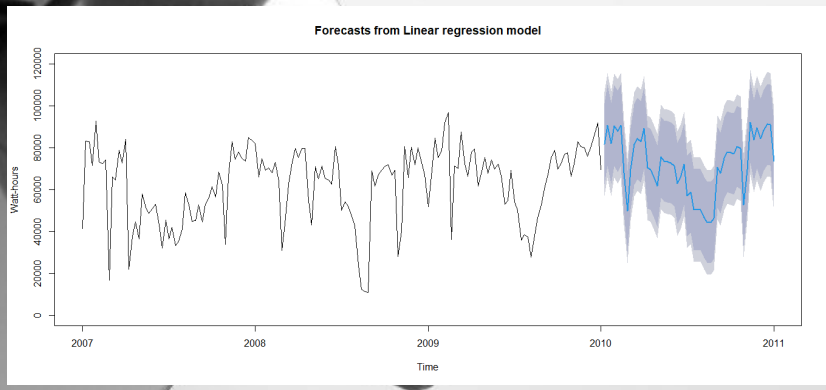

Figure 15: Linear Regression Forecast of Sub-Meter 3 Usage for 2010

- We can project the following year's anticipated consumption on each individual sub-meter using a linear regression model.
- Figures 13-15 display forecasts for 2010 consumption accounting for seasonality and established trends in usage.
- The model clearly anticipates a dip in summer consumption on all three sub-meters.
- Sub-meter 1 appears to maintain a level trajectory.
- Sub-meter 2 continues on its slight downward trend.
- Sub-meter 3 maintains its established seasonal volatility while continuing its upward trajectory.
- The figures indicate a range of potential usage with specificity to the time of year.

Data Decomposition

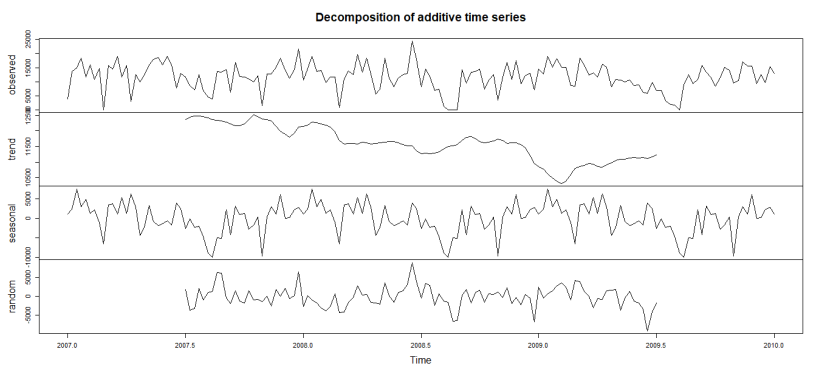

Figure 16: Decomposition of Sub-Meter 1

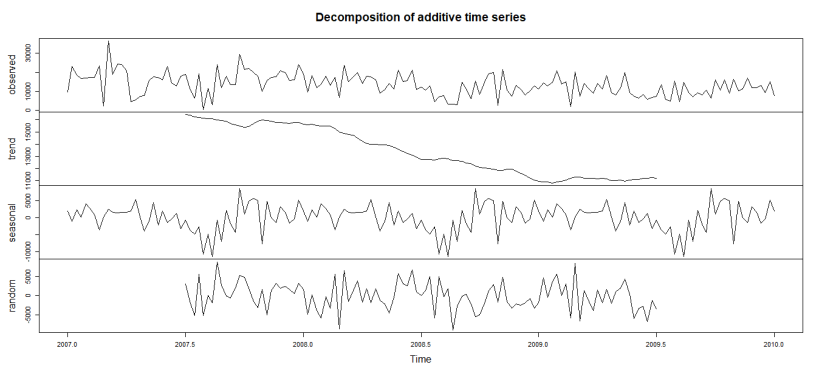

Figure 17: Decomposition of Sub-Meter 2

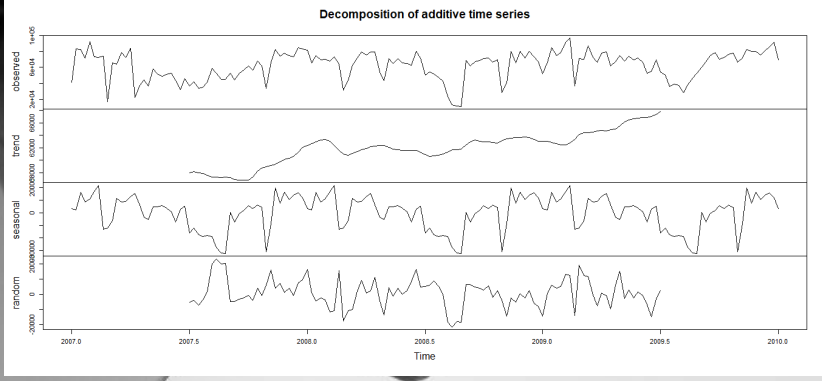

Figure 18: Decomposition of Sub-Meter 3

- Several factors are taken into account while trying to forecast future usage. These include historical trajectory (trend) and seasonality, as well as unforeseen (and unpredictable) factors.
- We can break down the data to display these various components through a modeling process known as decomposition.
- In so doing, we are able to analyze each element as it relates to the observed data.

	Sub-Meter 1			Sub-Meter 2			Sub-Meter 3		
	Seasonal	Trend	Random	Seasonal	Trend	Random	Seasonal	Trend	Random
Minimum	-9982.09	10315	-9161.71	-11331.4	10832	-8913.35	-32286.8	56863	-21644
1st Quartile	-2061.08	11163	-1661.08	-1711.2	11238	-2147.12	-7570.53	60806	-4242
Median	890.355	11614	-278.41	284.53	12823	-36.43	3415.67	62334	812
Mean	7.235	11599	-93.09	12.86	13384	-84.39	21.76	61977	740
3rd Quartile	2828.951	12141	1485.13	2206.77	15537	1978.34	9077.63	63335	5693
Maximum	7591.259	12503	8975.53	8628.73	16390	8744.57	21689.3	67748	23124

Holt-Winters Forecasting

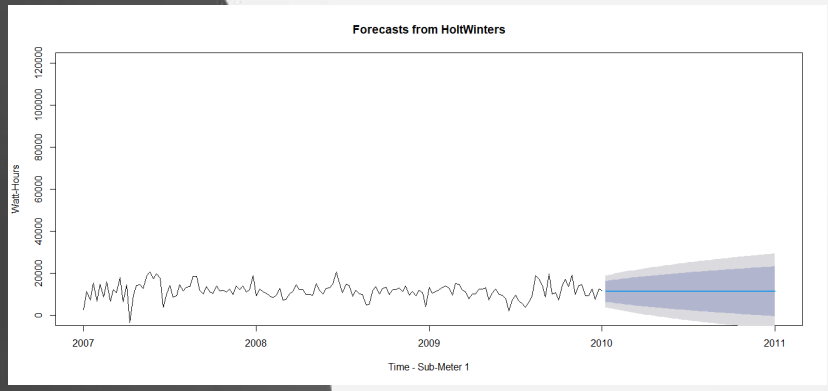

Figure 19: Holt-Winters Forecast for Sub-Meter 1

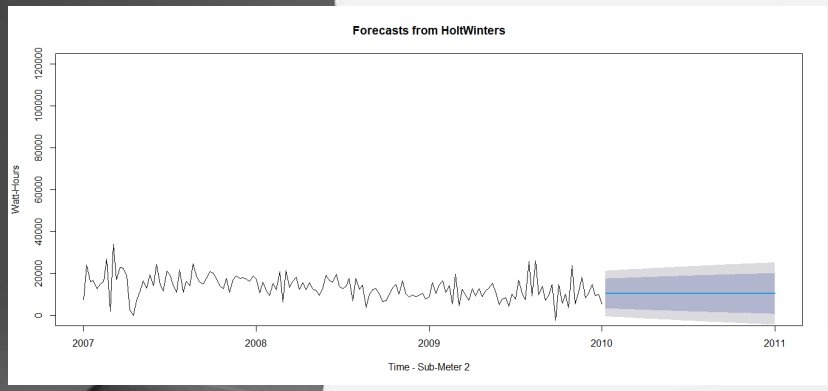

Figure 20: Holt-Winters Forecast for Sub-Meter 2

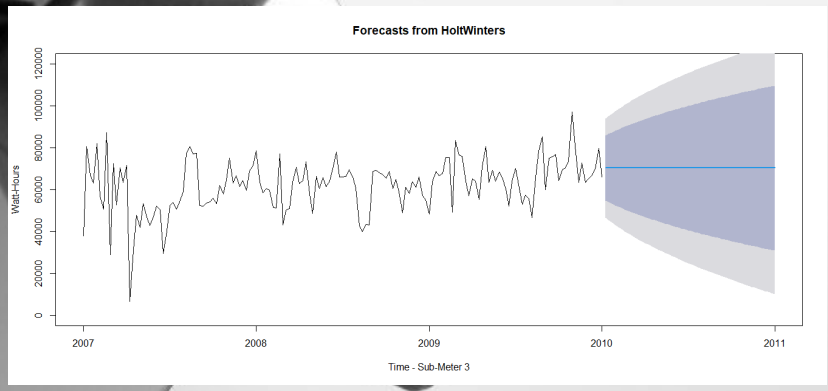

Figure 21: Holt-Winters Forecast for Sub-Meter 3

- By removing the seasonal element of the data, we can model a more general forecast of energy usage through the Holt-Winters method.
- This method smooths the data to project a reasonable range of anticipated usage without overfitting the historical data.
- The further we try to forecast, the less confident our predictions are.
- These forecasts can anticipate an average usage over the course of an extended period of time and could be used for "budget billing," where a customer is billed the same amount each month based on anticipated yearly usage.

Observations

- Sub-meter 3 uses by far the most energy and is the most seasonally affected.
- Summer time sees the lowest energy consumption, while winter sees the highest.
- Sub-meter 1 sees its most activity on weekends and evenings.
- Use of AC and/or water heater seems to be trending upward while manual appliances (washer/dryer, oven, dishwasher) are on a downward trend.
- Overall, energy consumption is highly correlated with time of day as well as time of year, indicating a strong influence by both the resident's daily routines as well as seasonal changes.

Business Recommendations

- It makes sense to keep appliances that draw energy on a predictable cycle (refrigerators, AC units) on separate sub-meters from those that are user activated (ovens, dishwashers) for efficient energy management. Being able to easily reference how much and how often energy is used on automatic cycles allows the user to program responsible energy settings.

- Installing smart thermostats that communicate with personal devices (smart phones) could reduce power consumption by allowing the system to know when the resident is home or away, thus scheduling AC use more responsibly.

- Providing residents with projected use figures could incentivize conservation of energy.

- Knowing that the AC and water heater are the most consistent power sucks, it would be wise to investigate energy efficient models to install in new units.

- If possible, charging higher rates for energy during peak usage times could incentivize residents to be more thoughtful in their energy consumption.

Lessons Learned

- Using raw time series data when there are so many data points is like trying to analyze an entire movie frame by frame. For higher level analysis, aggregating data is much more helpful.

- Linear regression modeling for time series analysis yields nice graphics that look like they fit, but given the random nature of real world data, may not be the most reliable. Holt-Winters provides less specific predictions but set more realistic expectations.

- My takeaway from this project is that it only scratched the surface of what time series analysis is really capable of.