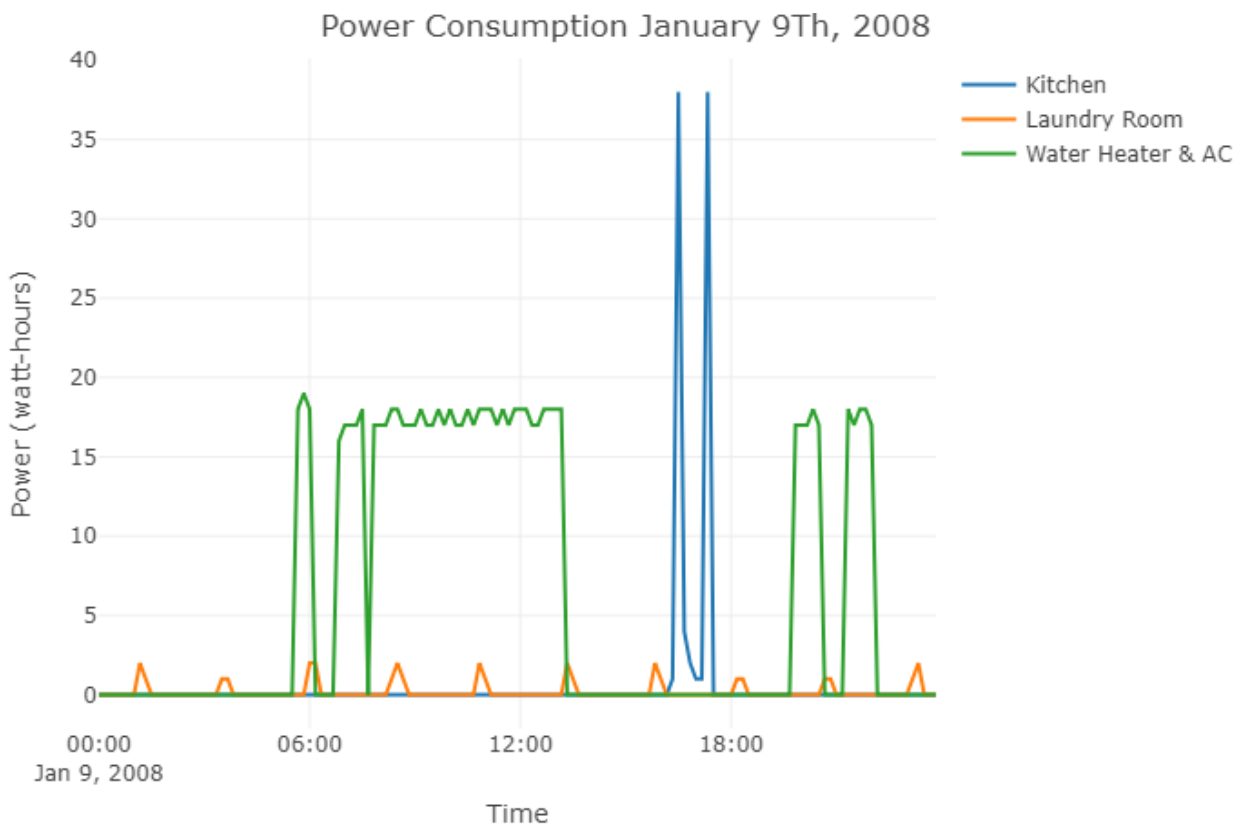
Smart Homes' Energy Data Analytics

Xtol: Deep Analytics and Visualization 2017.3, task 2

Author: Esteban Villalobos Gómez

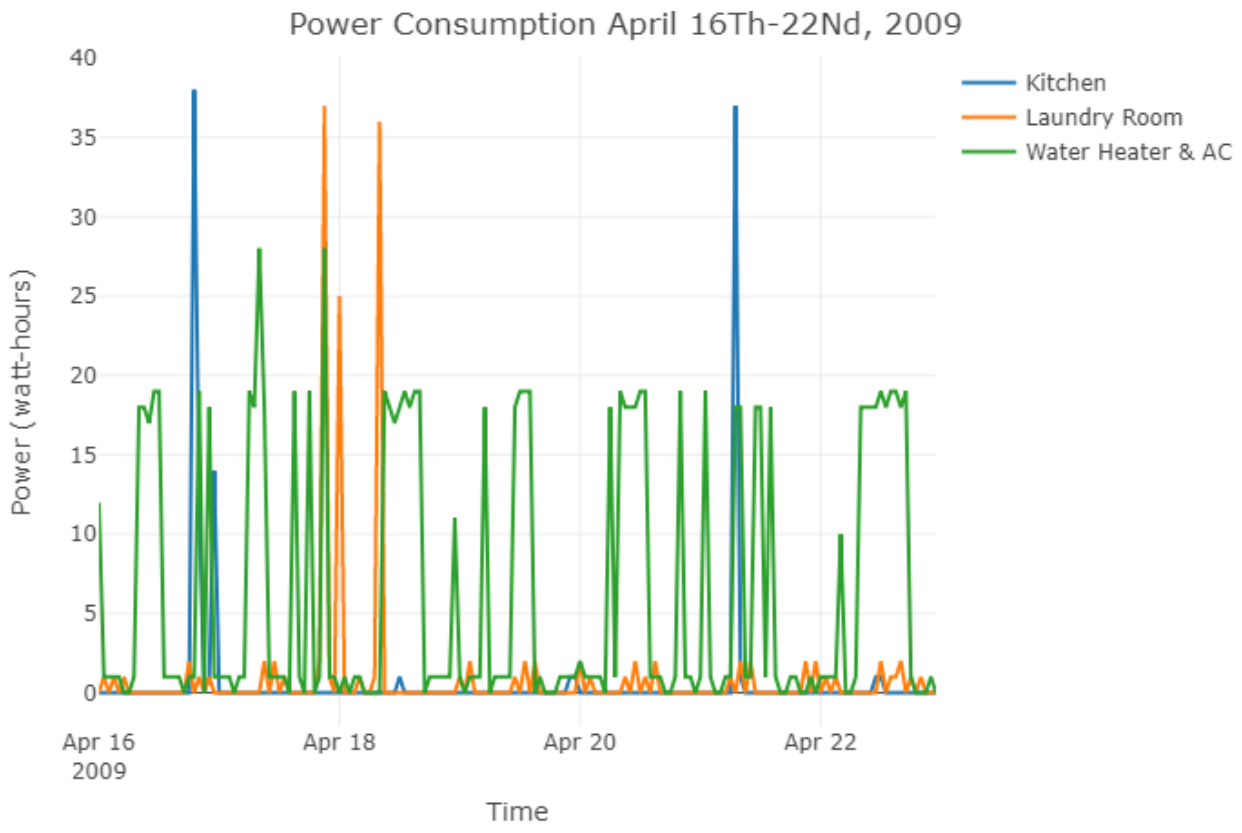
October 24th, 2019

Single day view



- This graph show customers the power consumption of their home on any single day
- Customers can see patterns easily, for instance, there is an appliance consuming power periodically in the Laundry Room, when not in use
- The graph shows samples taken every 10 minutes

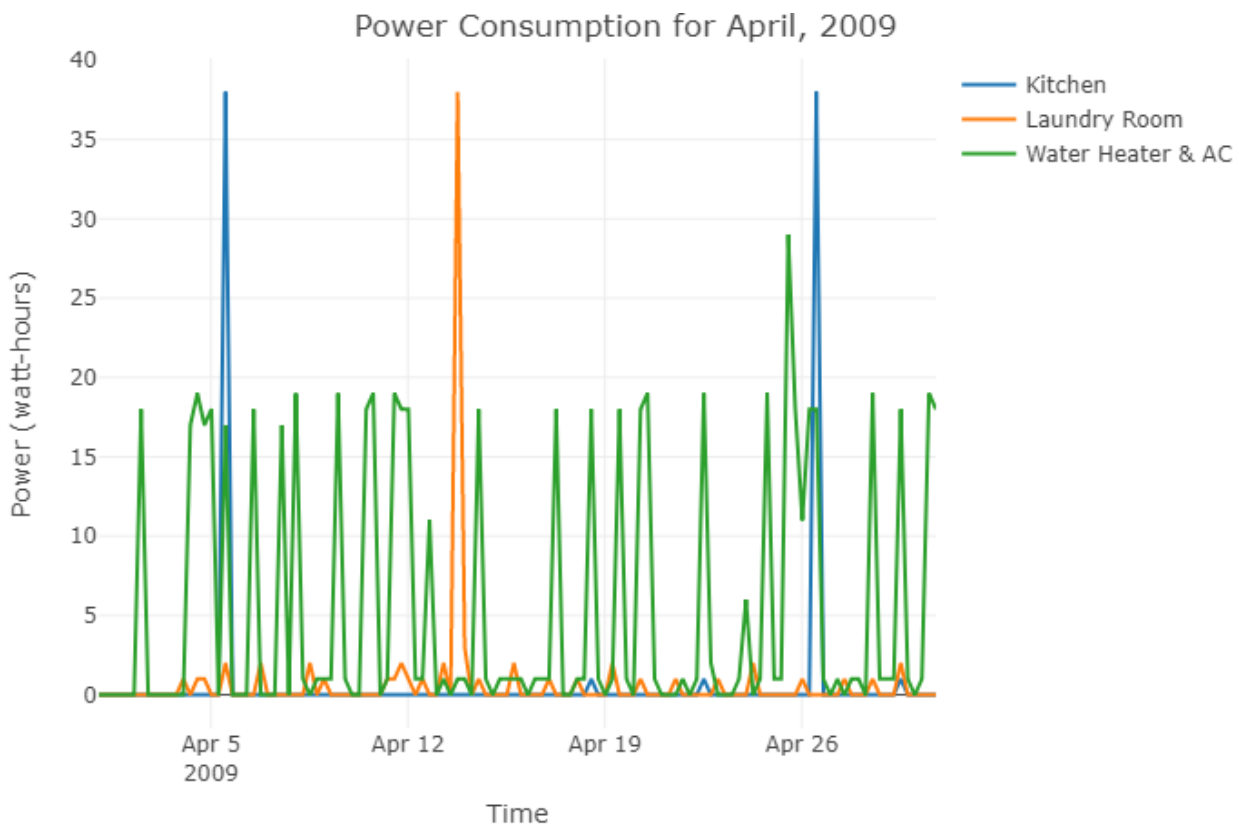
Week view



- This view does “zooms out” to show patterns of power consumption on several days
- It is noticeable how the Water Heater and the AC are constantly drawing power

  Perhaps the AC is not turned off when no one is home?
- Both laundry and cooking happen not very frequently
- Samples taken with a 60 minutes frequency

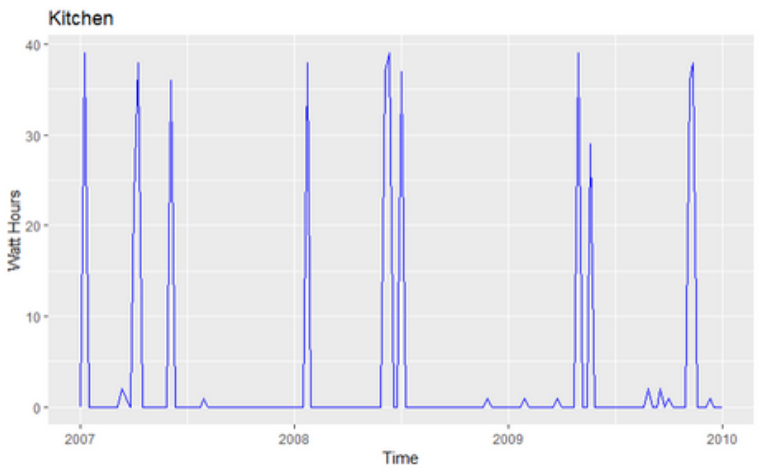
Month view

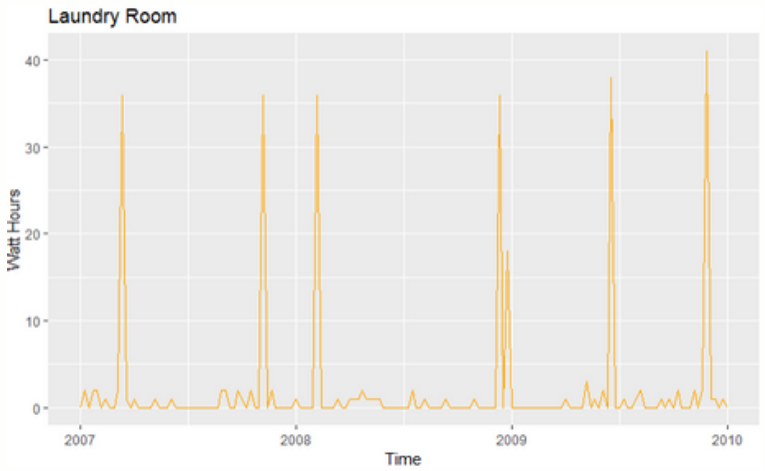


- This view “zooms out”even more, to show a month of power consumption
- There are similar patterns in a month to the weekly view, which reinforces that the family has very well established routines
- It would be interesting to correlate the Water Heater and AC usage with the outside temperature.
- Samples taken with a 6 hour frequency

  Be aware that depending the frequency of the samples, some data points might be hidden

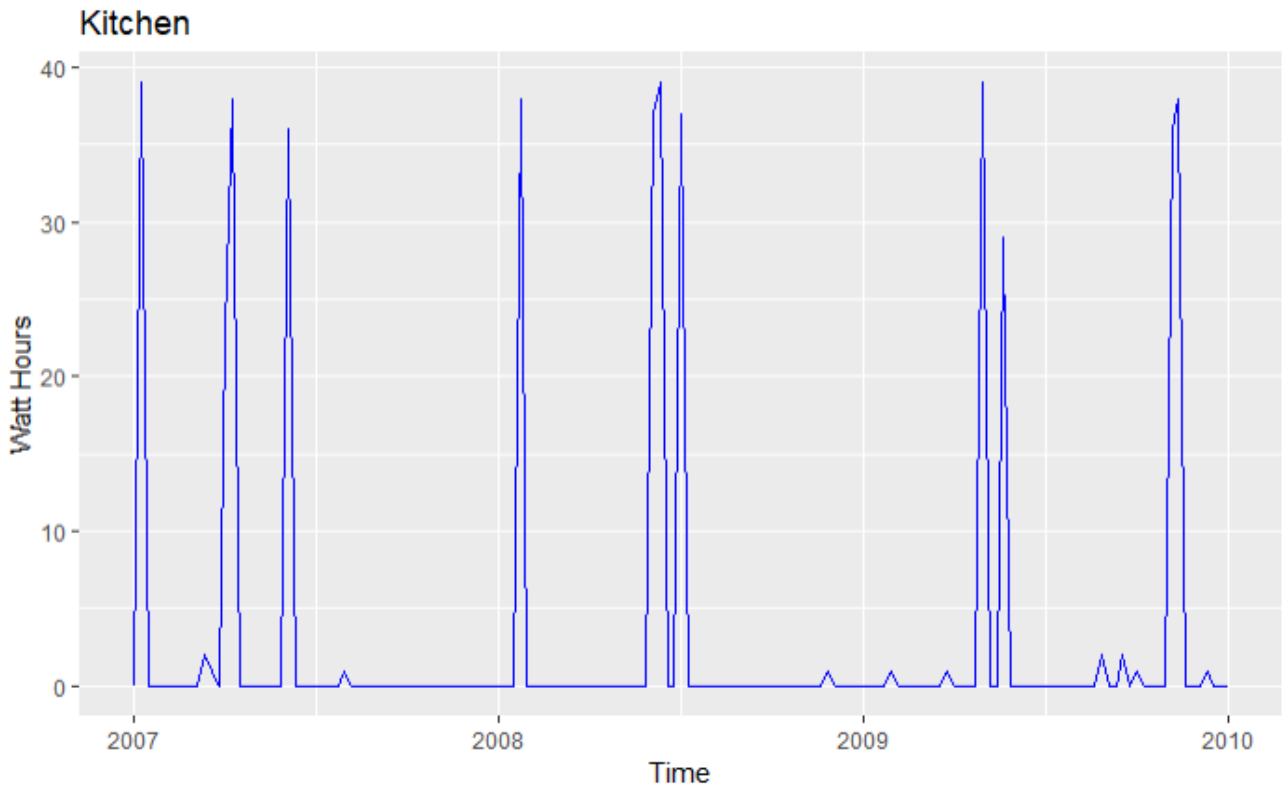
All three sub-meters time series over 3 years





Kitchen 3Y time series

- The time series indicates a seasonal use of the kitchen appliances that are electrical.



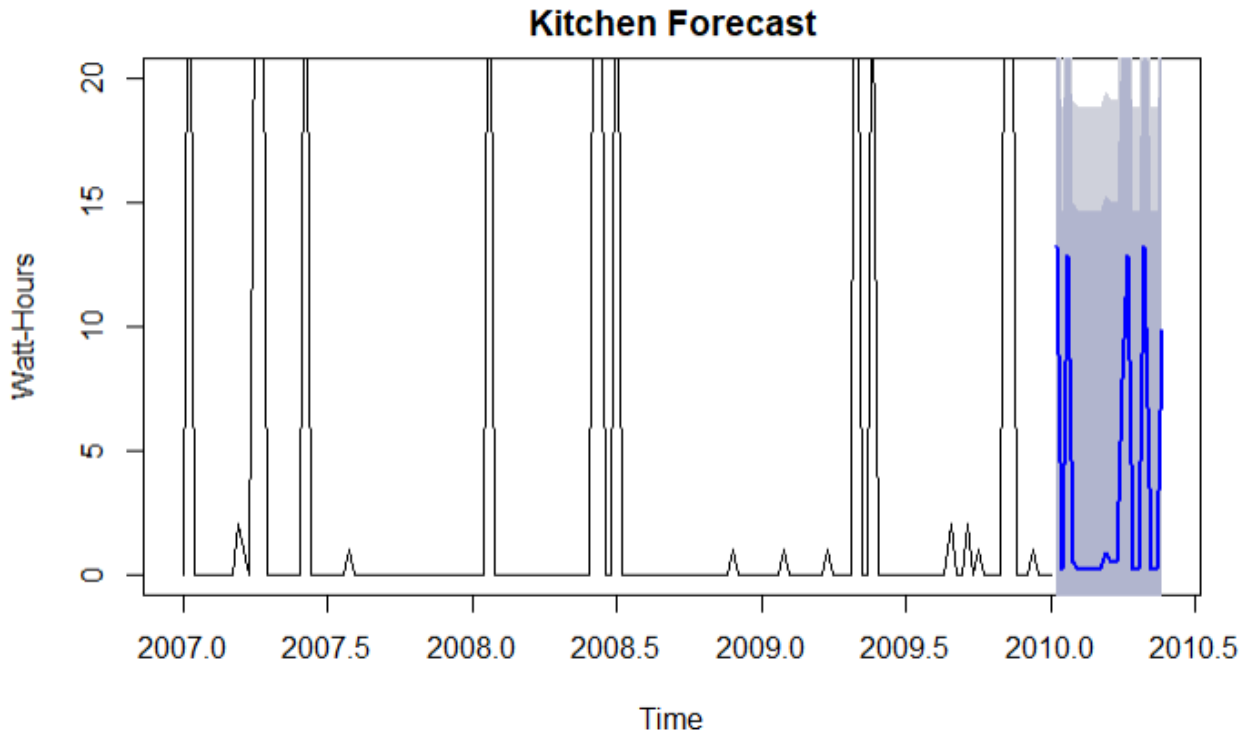
Laundry Room 3Y time series

- The time series indicates as well a seasonal use of the laundry room, specially during wet/winter months

Water Heather & AC 3Y time series

- There has been a constant usage of both the Water Heater and the AC over the years
- Most of the electricity bill is derived from these appliances

Linear Regression Forecast

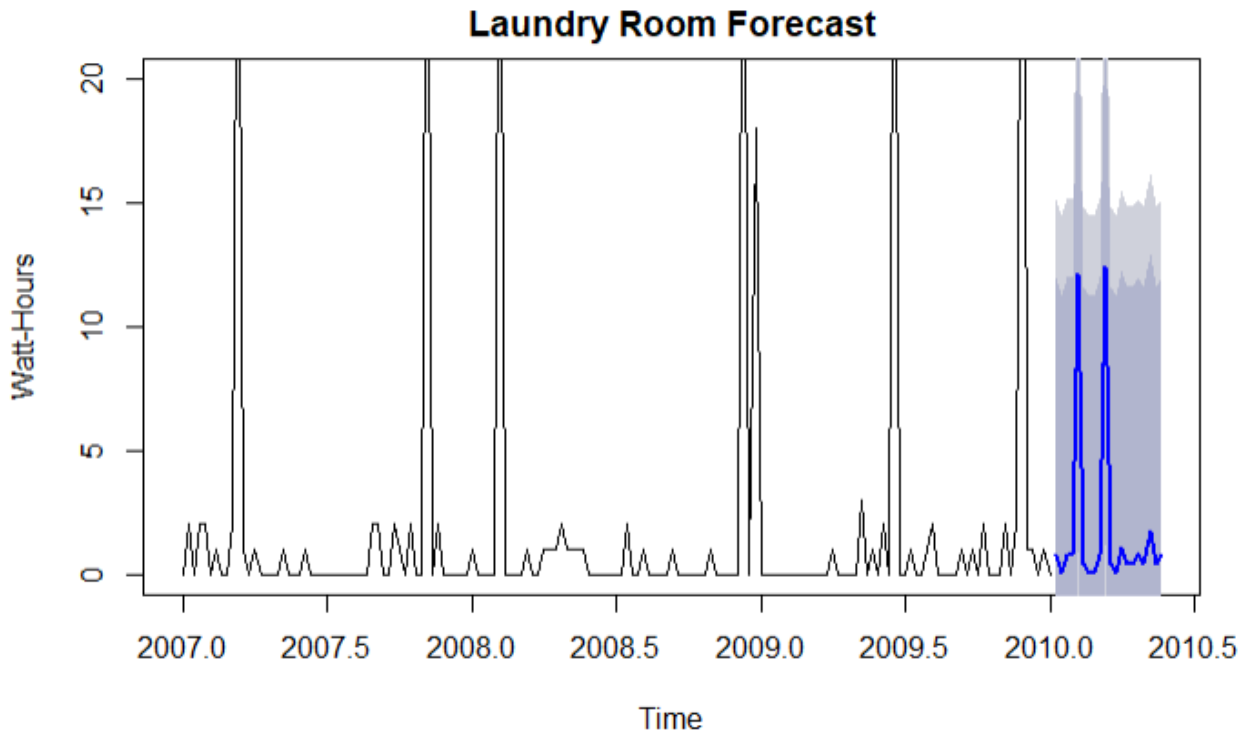


- Using a simple prediction model a customer can predict it's power consumption
- The gray shadows point the confidence range for the prediction

  light-gray = 90% confidence

  dark gray = 80% confidence

Linear Regression Forecast

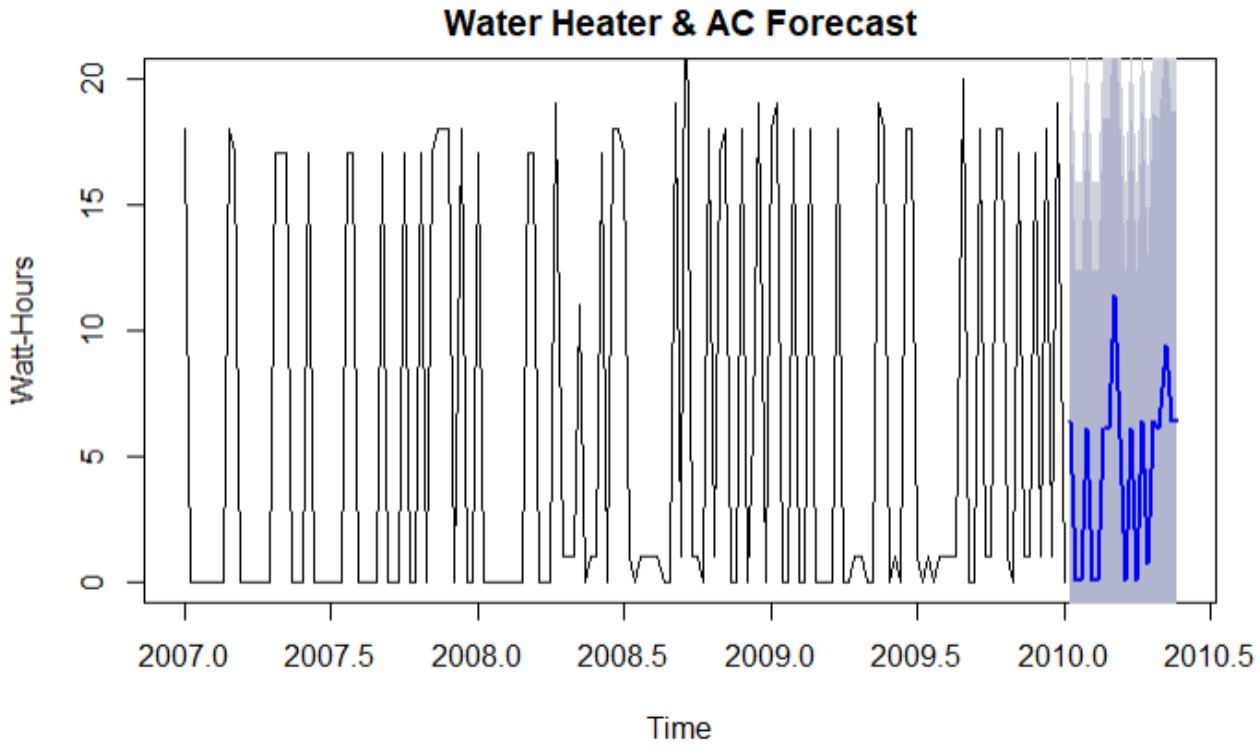


- The linear model can also predict the power usage of the laundry room, with the same levels of confidence

  light-gray = 90% confidence

  dark gray = 80% confidence

Linear Regression Forecast



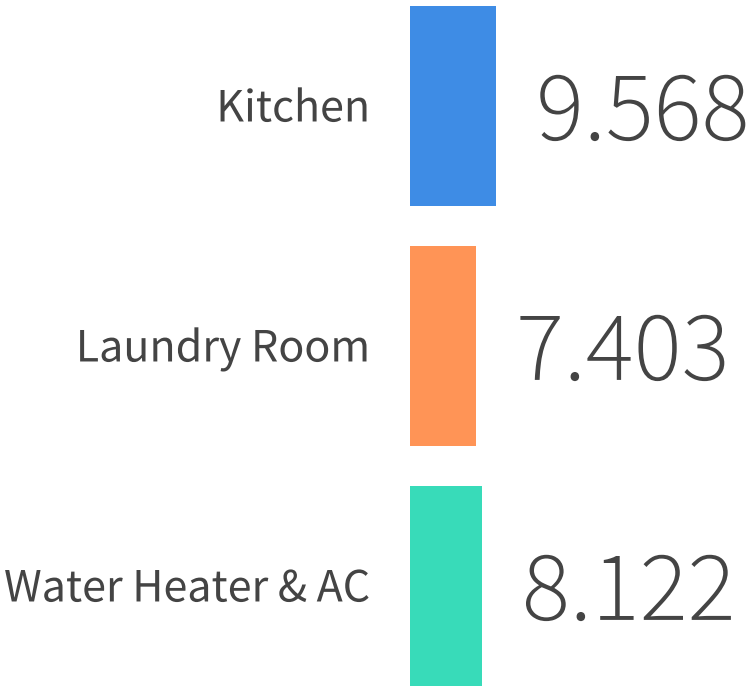
- In this case, the shadowed area of 80% and 90% of confidence, might point to a more realistic scenario, given the constant usage over the years

  light-gray = 90% confidence

  dark gray = 80% confidence

Prediction Performance

RSME metric



The RSME (Root Square Mean Error) measure is a common metric to validate each prediction model's accuracy. In this case all three models show quite low error rates

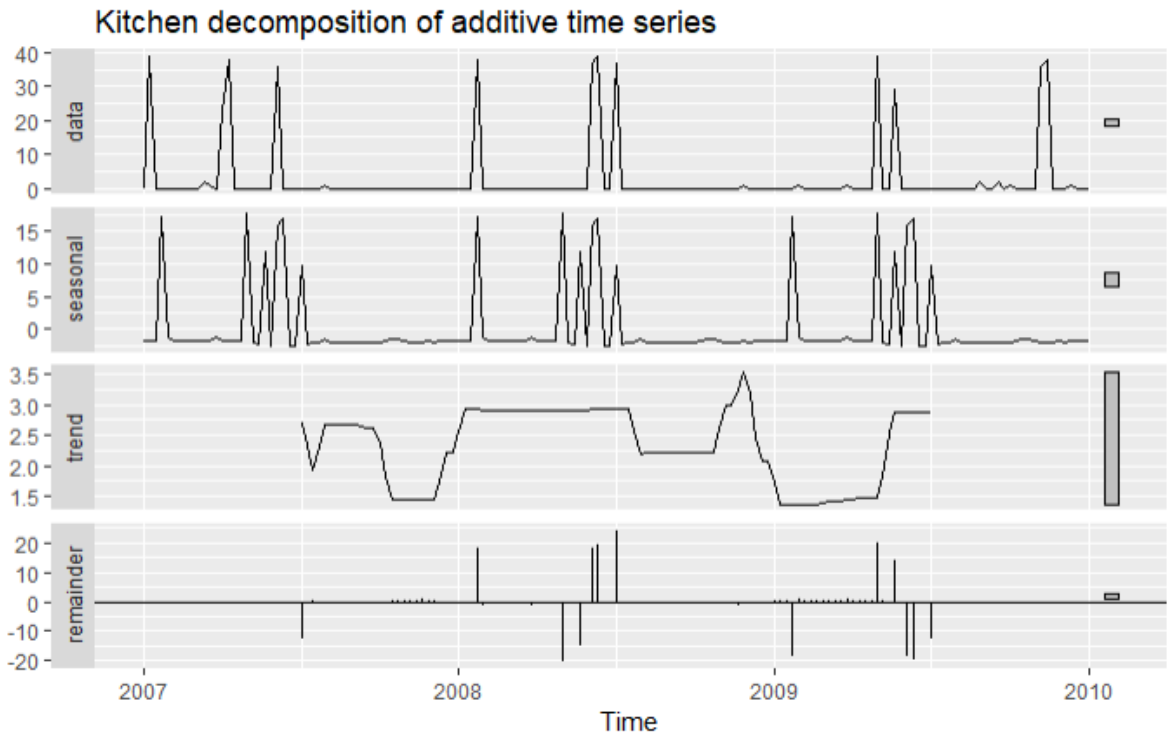
Prediction Performance

RSquared metric

0.3375	0.3375	0.3544
Kitchen	Laundry Room	Water Heater & AC

The R squared error measurement is also useful when evaluating prediction models, in this case is near zero, which suggest the models could be performing pretty well with the existing data, but might not react as well with new data

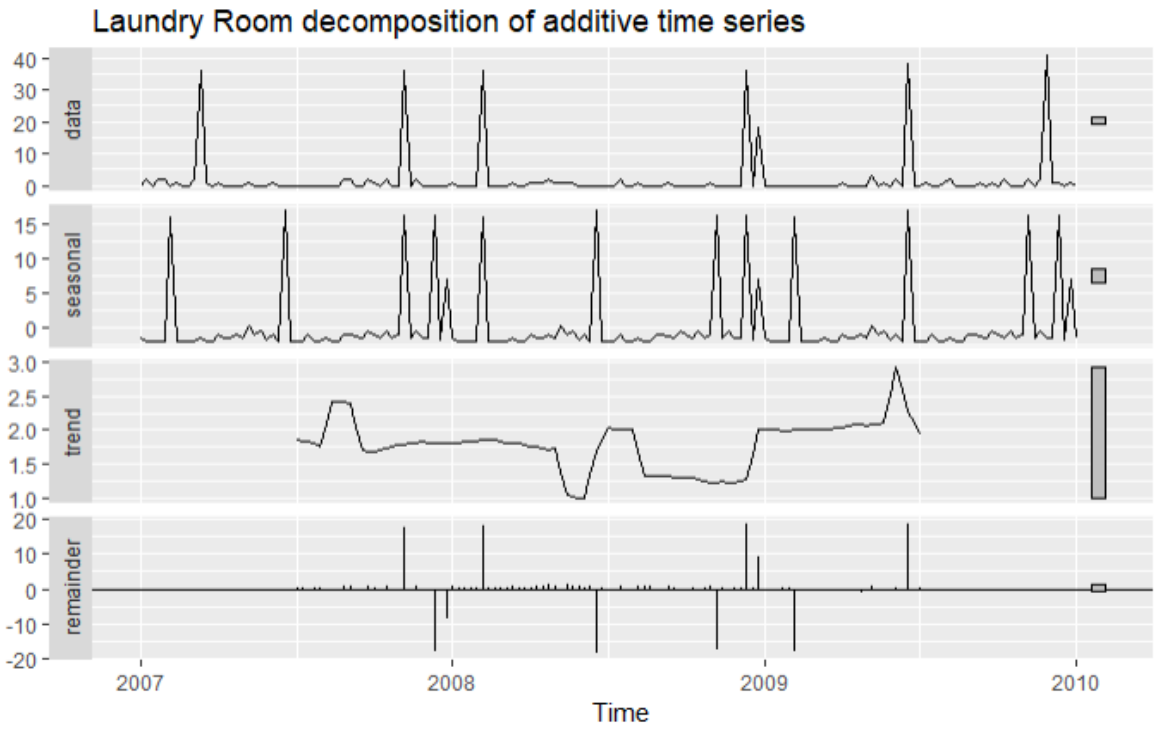
Time Series Decomposition



From the graph we can validate there is a seasonal effect on the power consumption

There is a more ore less stable trend on the Kitchen power consumption over time

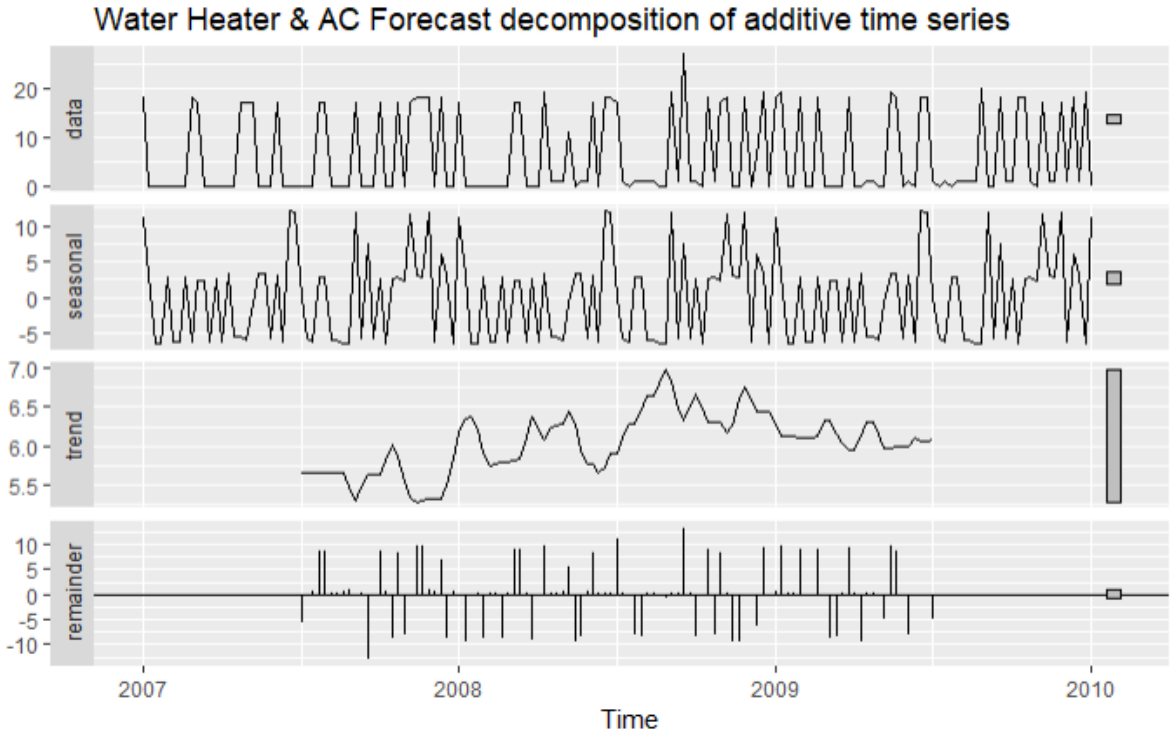
Time Series Decomposition



Laundry Room also seems to have some impact from the seasonal component

There is no an upward or downward tendency regarding the laundry room usage

Time Series Decomposition

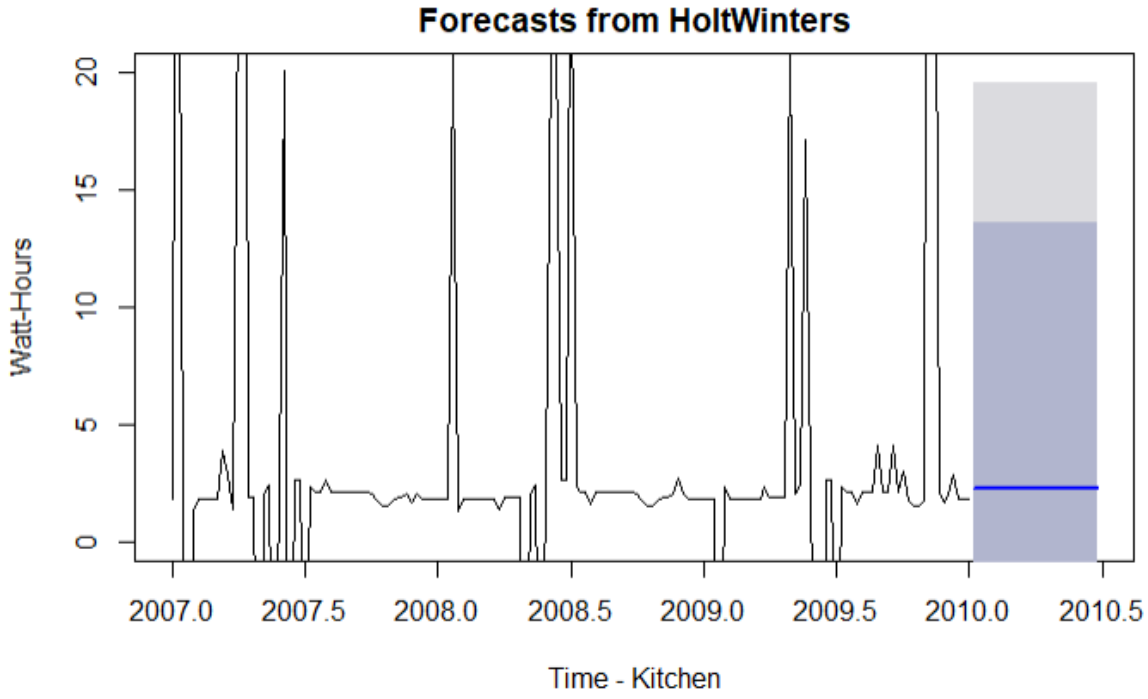


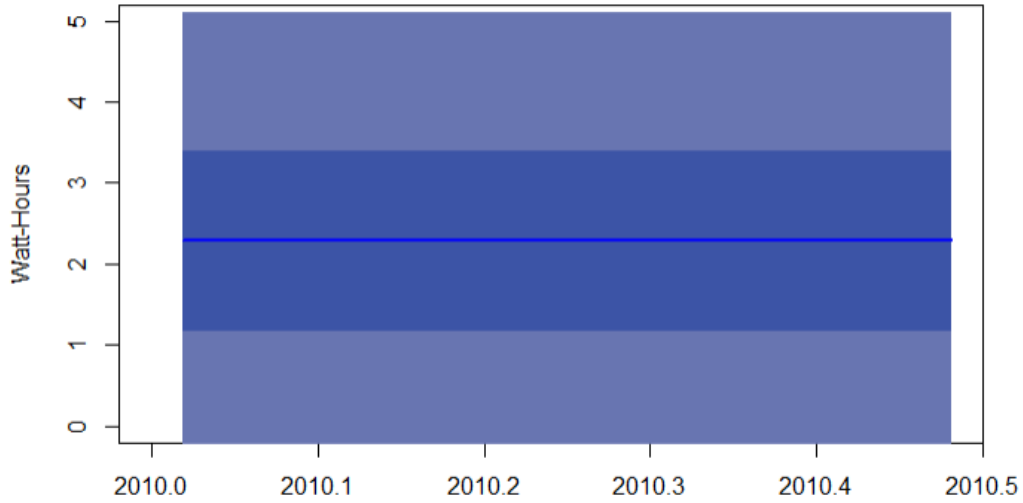
In this case the seasonal effect seems to have a bigger impact

There is an upward trend on the power usage, it might be related to higher changes in temperature over the years

Time series decomposition statistics (median)

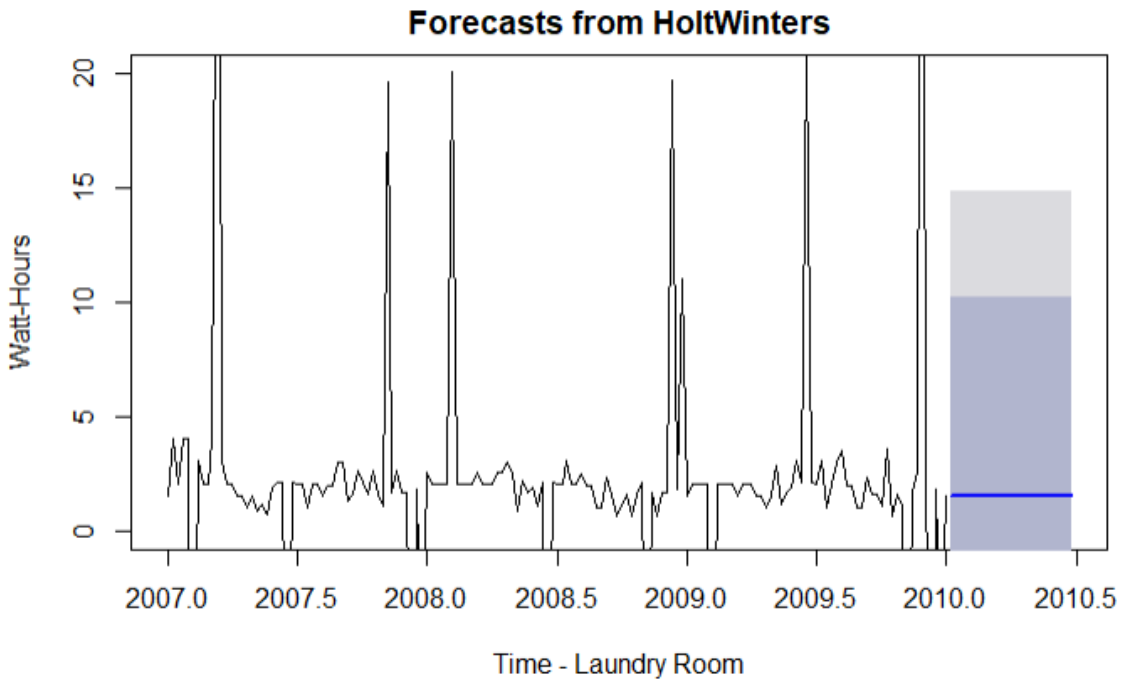
Kitchen Holt-Winters Forecast

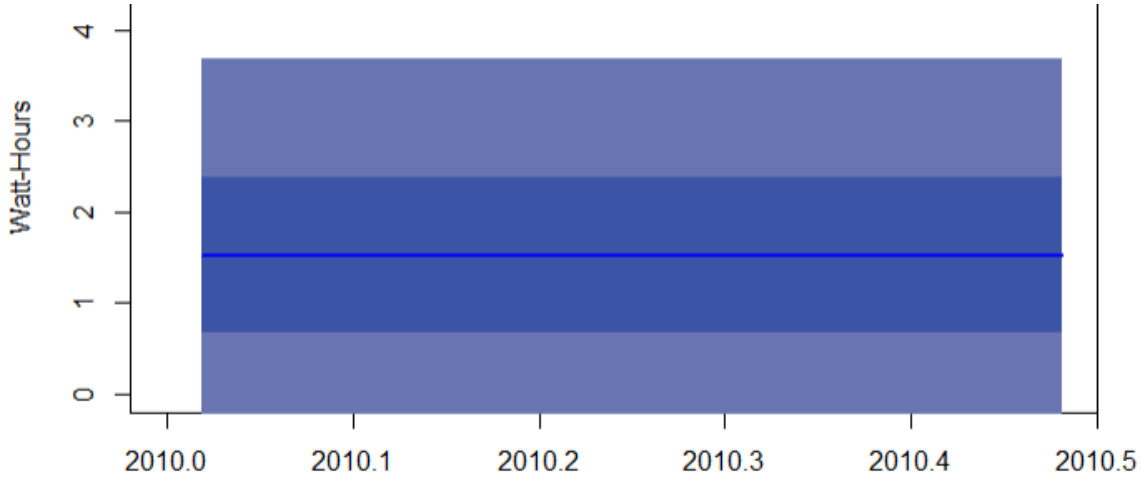




- This forecast takes into consideration the seasonal data
- Given the low usage of the kitchen, it is predicted almost as a constant value of **over 2 watts/hour** for the future.

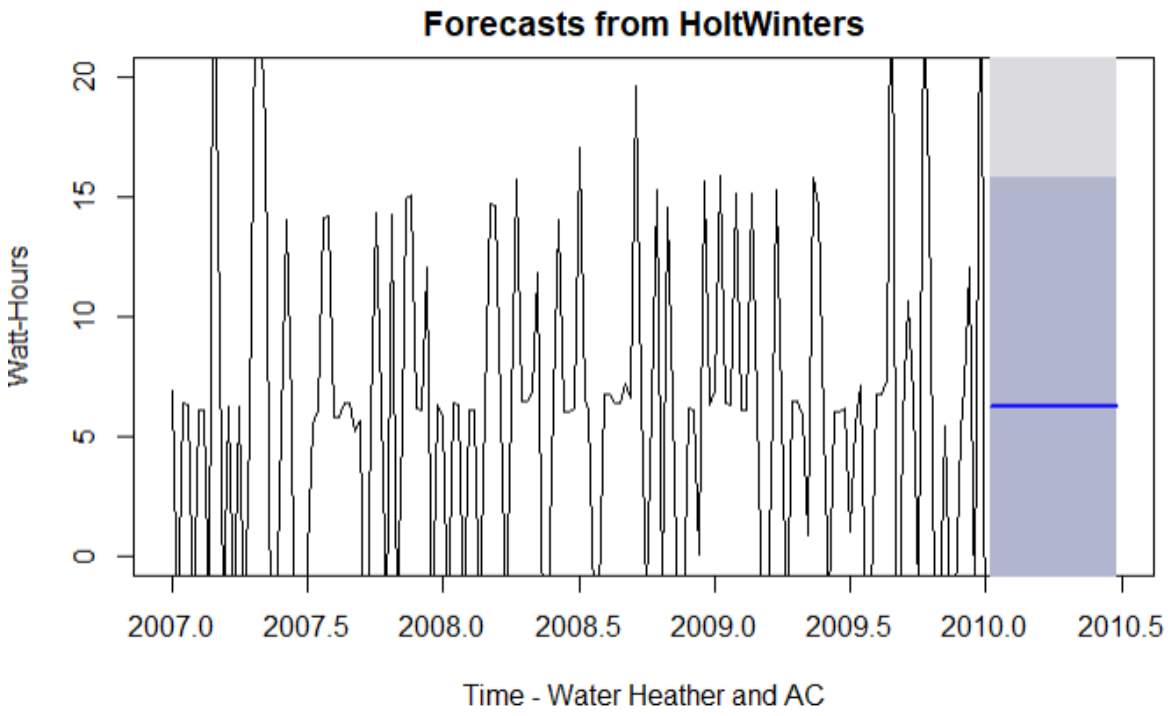
Laundry Room Holt-Winters Forecast

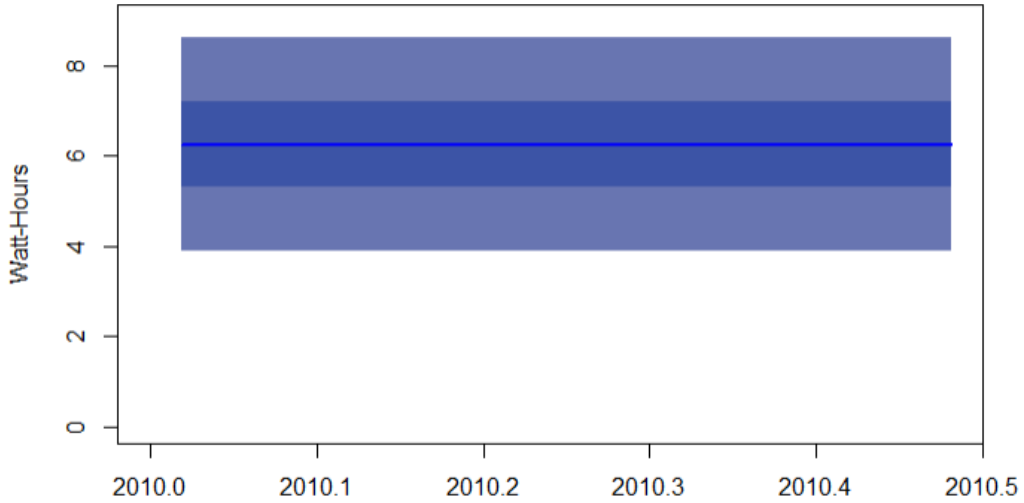




- Given the low usage of the laundry room, the forecast is a constant value of **less than 2 watts/hour** after removing the seasonal effect

Water Heater & AC Holt-Winters Forecast





- Given the constant use of these appliances the forecast is around **6 watts/hour** being the highest forecast of the three

Forecasts Improvements

- The time series collected can be correlated with other data sources that affect people's living patterns
- Taking into account outside temperature, season, and holidays would be a good insight for improving forecasts
- Adding more information about the number of family members, as long with their age would also add useful data to correlate with the sub-meters measurements.

Goal

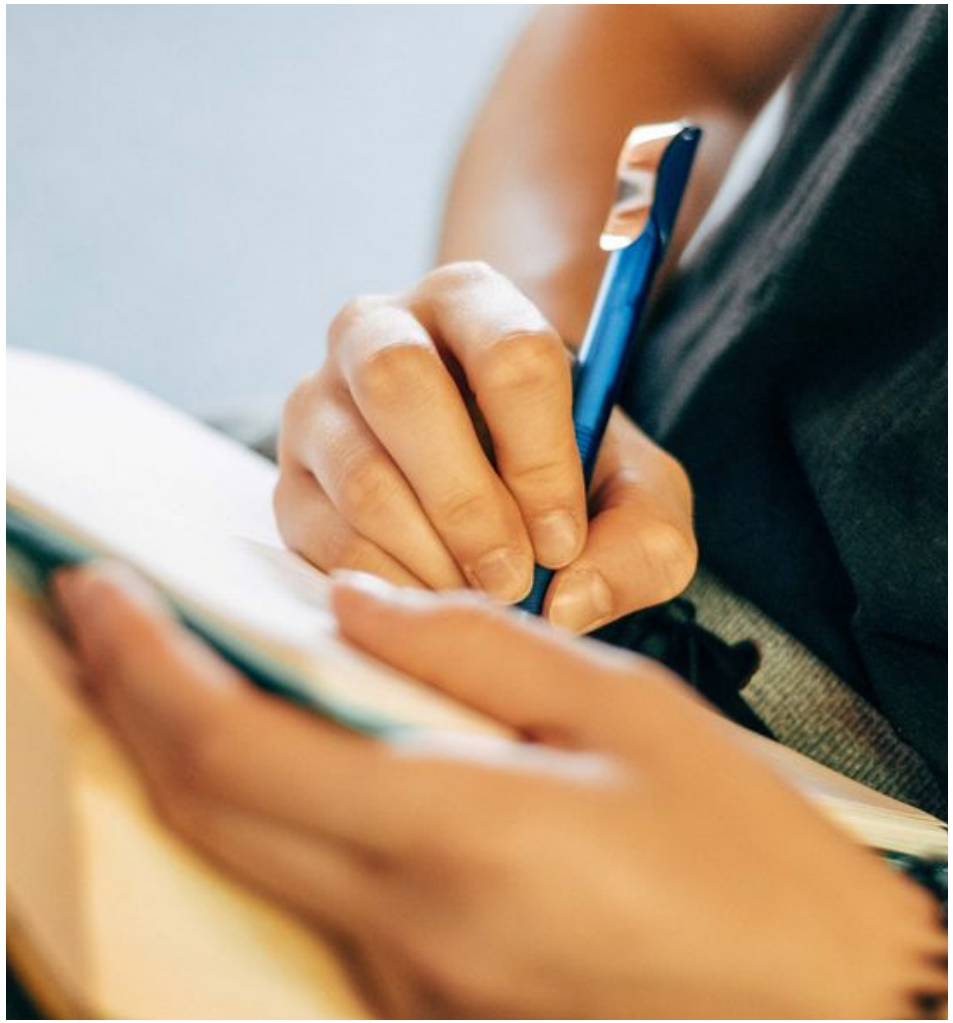
“Explore by visualizing the data and build predictive models to the home owners make decisions about altering power consumption”

- The visualizations and predictions shown in this report will help determine the areas of more power consumption for home owners

Business recommendations

1. Locate sub-meters in areas of the home where makes more sense: Living rooms, kitchen, laundry rooms, and also monitor big appliances individually (AC, Water Heater)
2. Correlate data gather from sub-meters with weather and holiday data
3. Collect if possible the number of family members of the household, and use it to provide more accurate reports
4. Create mobile apps with dashboards that home owners can easily use to see their current and past consumption as well as forecasts
5. The mobile app could be integrated with smart alarms, to determine when no one is home, and monitor the “passive power consumption” to alert home owners of possible “energy leaks” home

Lessons learned



- When zooming into data, given that some appliances don last more than an hour under use, depending on the specific minute we choose, we will see or not some important data points, like the high peaks consumption instants.

- Too many values of zero in the measurements make some statistics deceiving (like the mean) and also could cause some issues when predicting.

- Do not under estimate the simplest models (like linear regressions), in this case they gave excellent results

Thank you!