



# Energy Data Visualize and Analyze

---

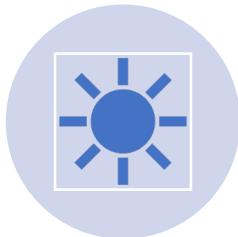
Sugi Devarajan – C4T2

# Objective

---



Answer the occupancy question using the power consumption data of a house in Europe.



Was the typical true for summer of 2008?



Find the typical patterns of usage of this house



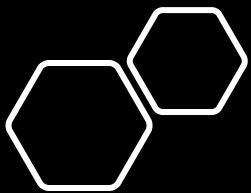
Any outliers to determine the claim?

▶ yr_2006	21992 obs. of 5 variables
▶ yr_2007	521669 obs. of 5 variables
▶ yr_2008	526905 obs. of 5 variables
▶ yr_2009	521320 obs. of 5 variables
▶ yr_2010	457394 obs. of 5 variables

# Data

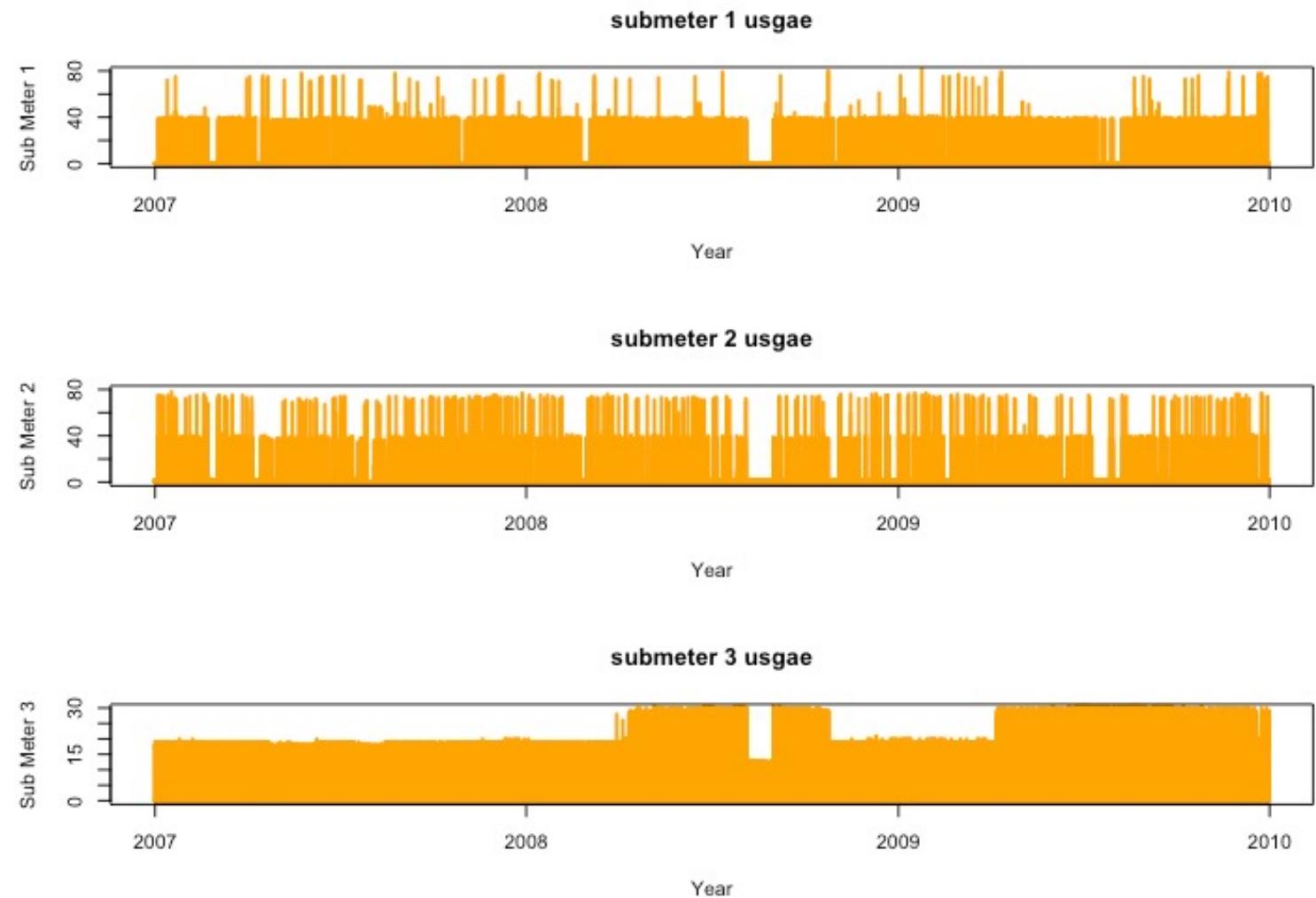
- Attribute Information:

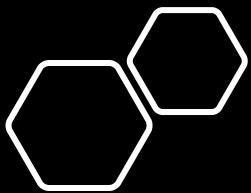
1. **Date**: Date in format dd/mm/yyyy
2. **Time**: time in format hh:mm:ss
3. **Global\_active\_power**: household global minute-averaged active power (in kilowatt)
4. **Global\_reactive\_power**: household global minute-averaged reactive power (in kilowatt)
5. **Goltage**: minute-averaged voltage (in volt)
6. **Global\_intensity**: household global minute-averaged current intensity (in ampere)
7. **Sub\_metering\_1**: energy sub-metering No. 1 (in watt-hour of active energy). It corresponds to the kitchen, containing mainly a dishwasher, an oven and a microwave (hot plates are not electric but gas powered).
8. **Sub\_metering\_2**: energy sub-metering No. 2 (in watt-hour of active energy). It corresponds to the laundry room, containing a washing-machine, a tumble-drier, a refrigerator and a light.
9. **Sub\_metering\_3**: energy sub-metering No. 3 (in watt-hour of active energy). It corresponds to an electric water-heater and an air-conditioner.



# Submeters Usage by Year

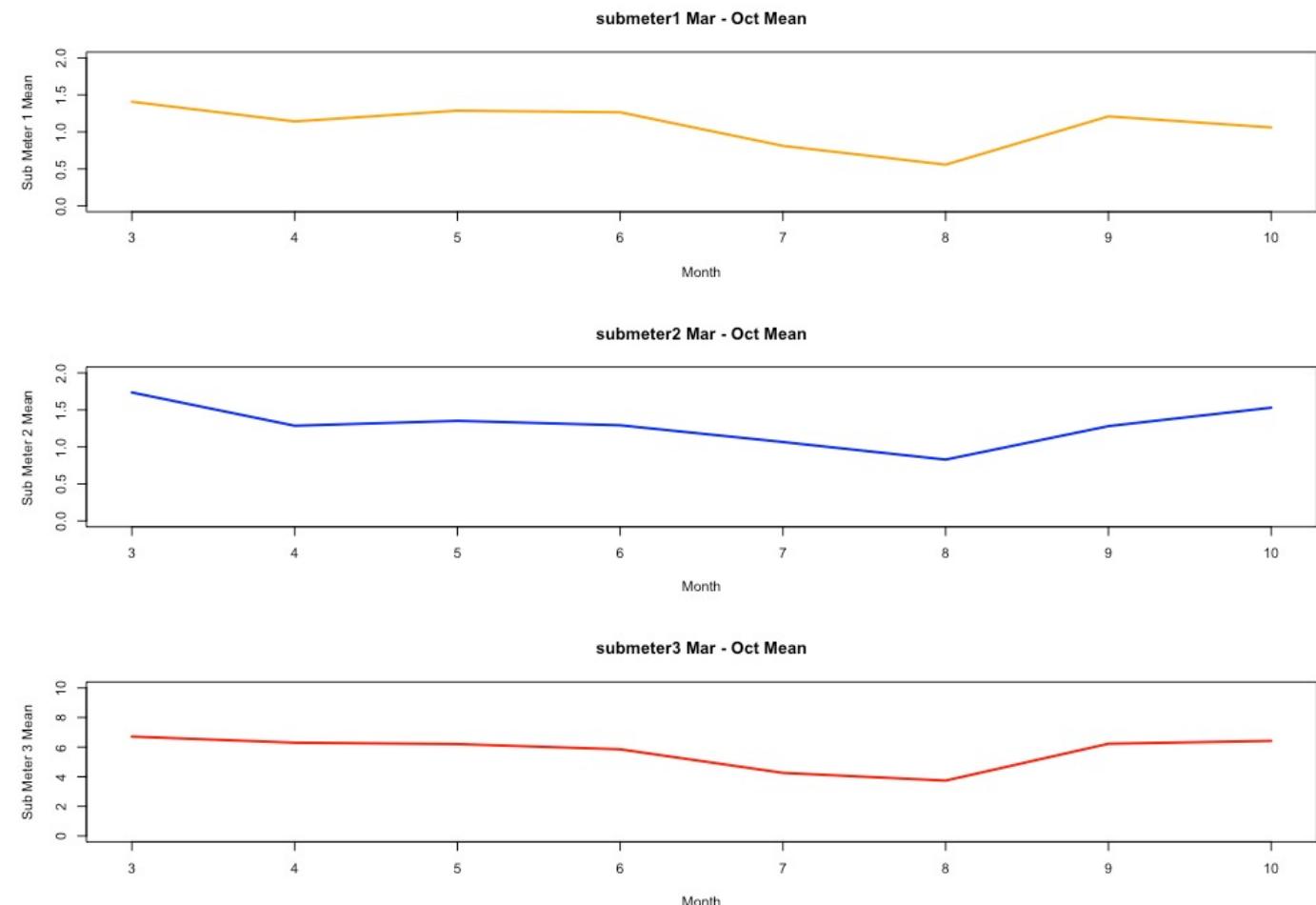
- Submeters usage for the 3 complete years we have in the dataset.

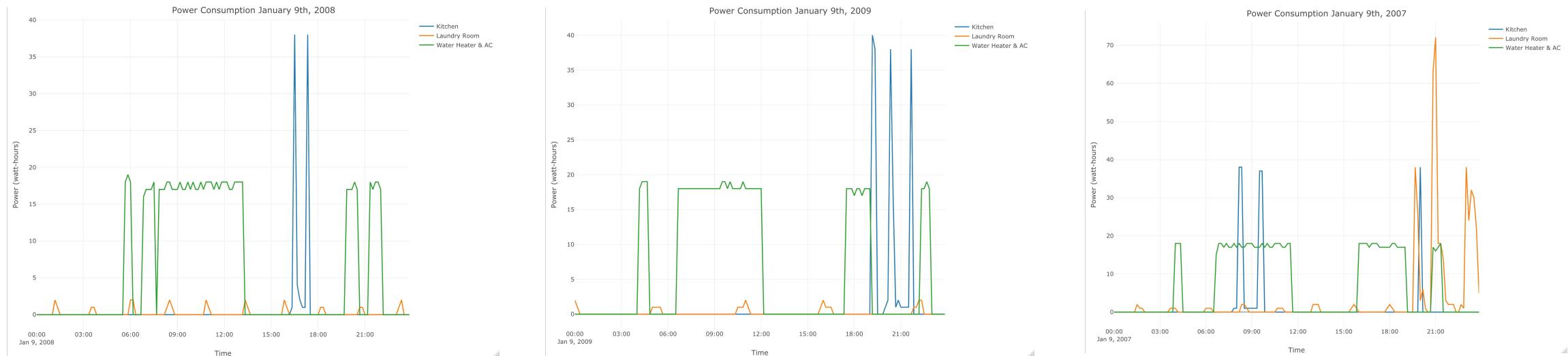




# Submeters mean graph

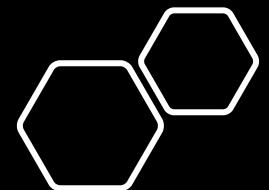
- Submeters mean is compared between march and October months. You can clearly see a dip in the July and August month.

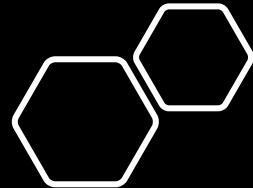




## Typical Patterns: Same Day Different Year

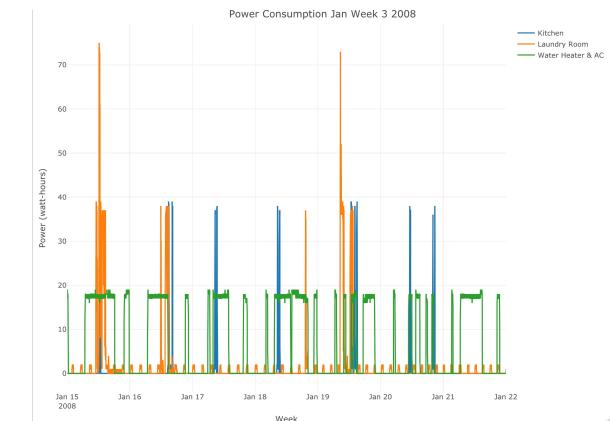
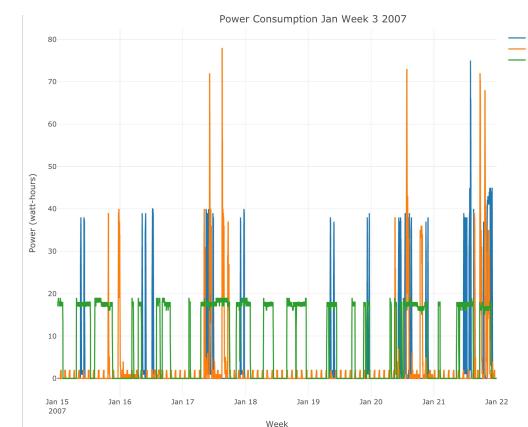
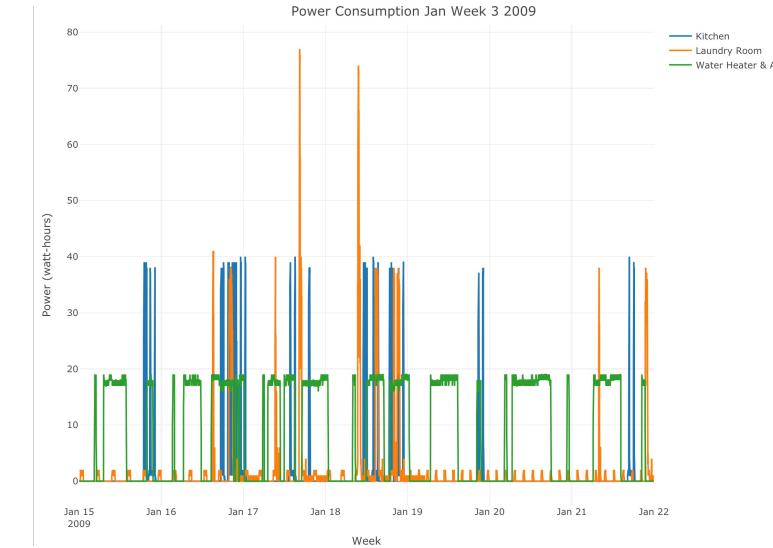
- Around 6 am they use the water heater
- From 7 am to 1 pm their AC is used mostly.
- They cook mostly in the evening.
- They use the water or AC at 9 pm.
- Their fridge is constantly running in intervals not using much power.
- They do Laundry very less but when they do it consumes the most power.

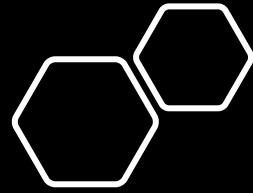




# Typical Patterns: Same Week Different Year

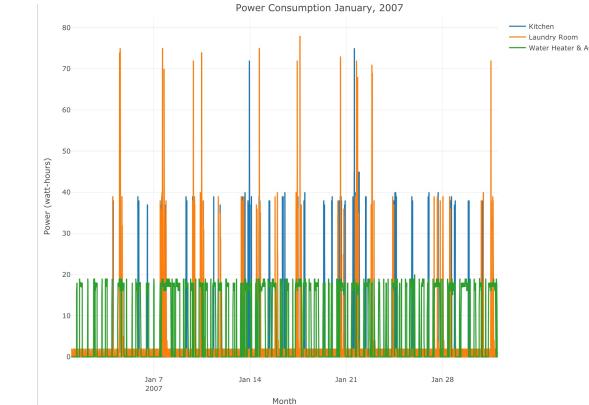
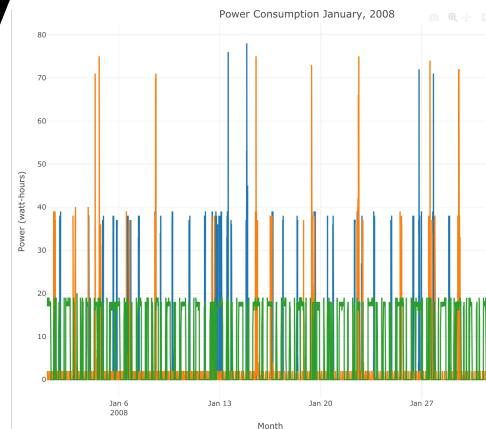
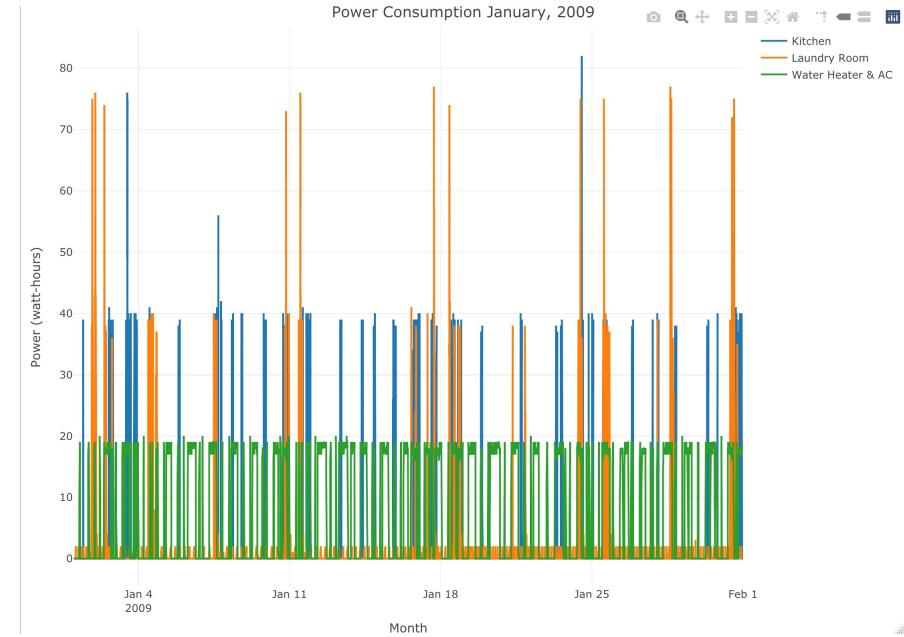
- They do Laundry twice or thrice a week.
- Cook at least 6 times in a week
- Use the AC and Water Heater everyday
- Fridge runs everyday

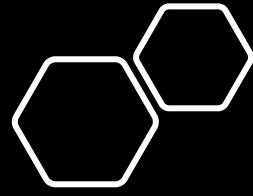




# Typical Patterns: Same Month Different Year

- Their drier must be consuming very high power and they should be doing heavy laundry 9 to 10 times in a month.
- Their AC and water heater are always use constant power.
- Twice or Thrice, they use their oven which consumes more power.

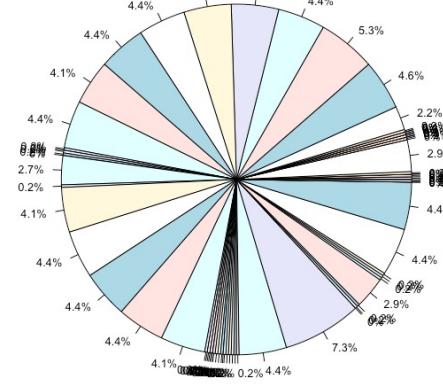




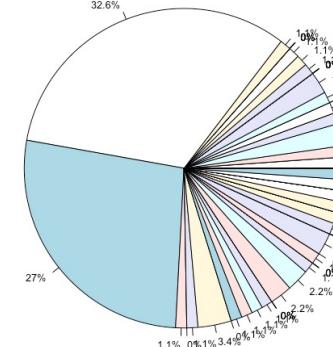
# Total Usage % at Various time of a day

- Submeter 3 runs constantly throughout day
- Submeter 1 is used more at two times in a day.
- Submeter 2 is used more twice in a day.

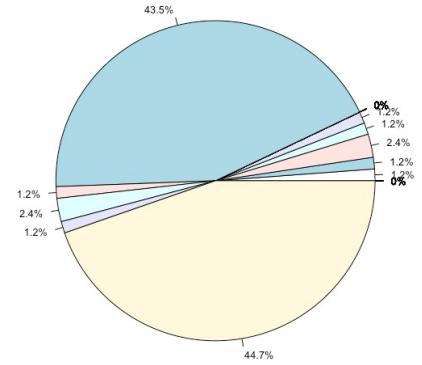
Sub meter 3 Piechart -Total Use at Various time of a day

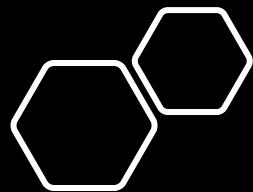


Sub meter 2 Piechart -Total Use at Various time of a day



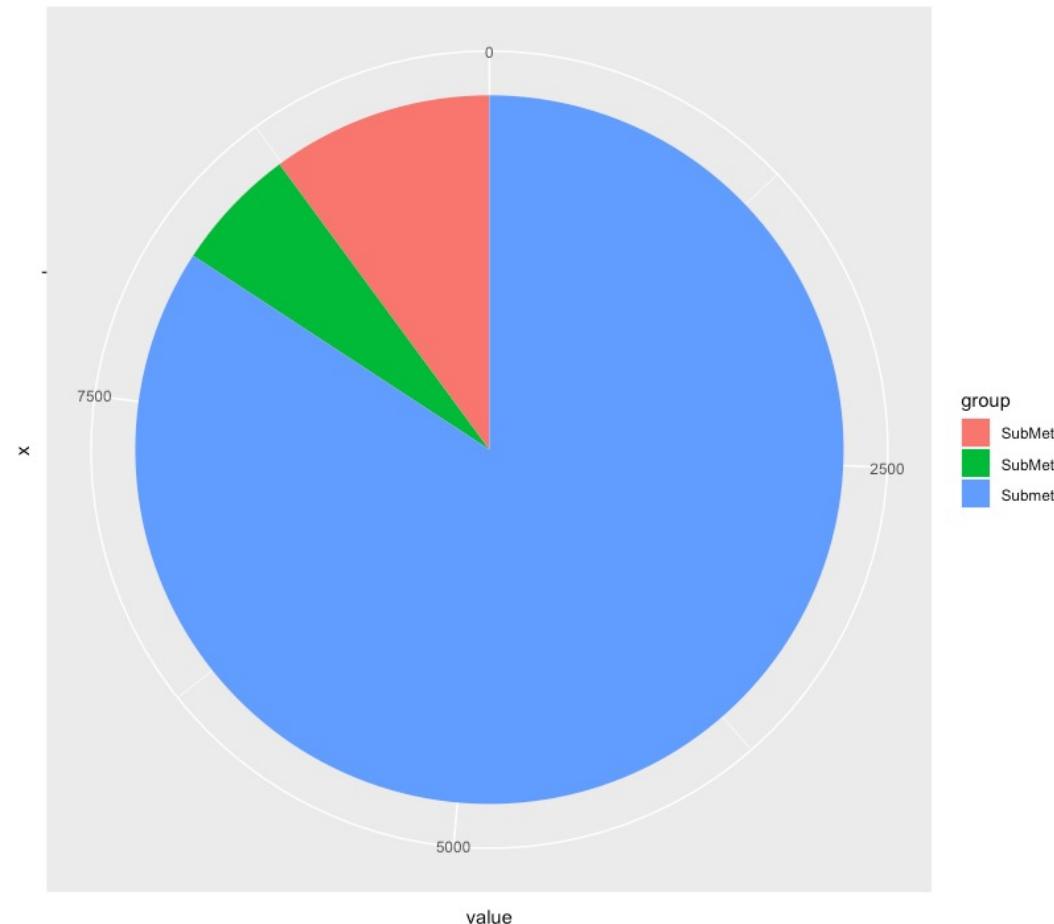
Sub meter 1 Piechart -Total Use at Various time of a day

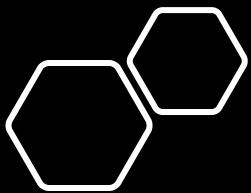




## Total power use over a day by each sub-meter

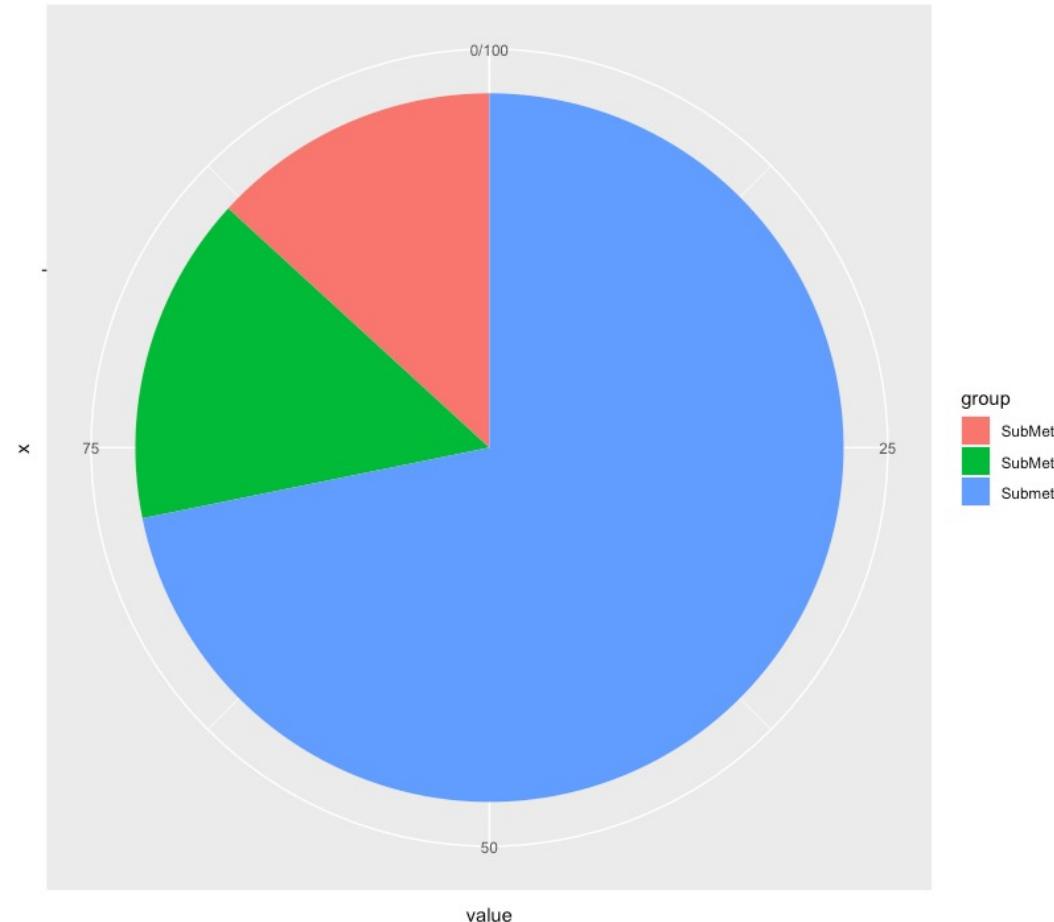
- Submeter 3 is used more than 1 and 2.





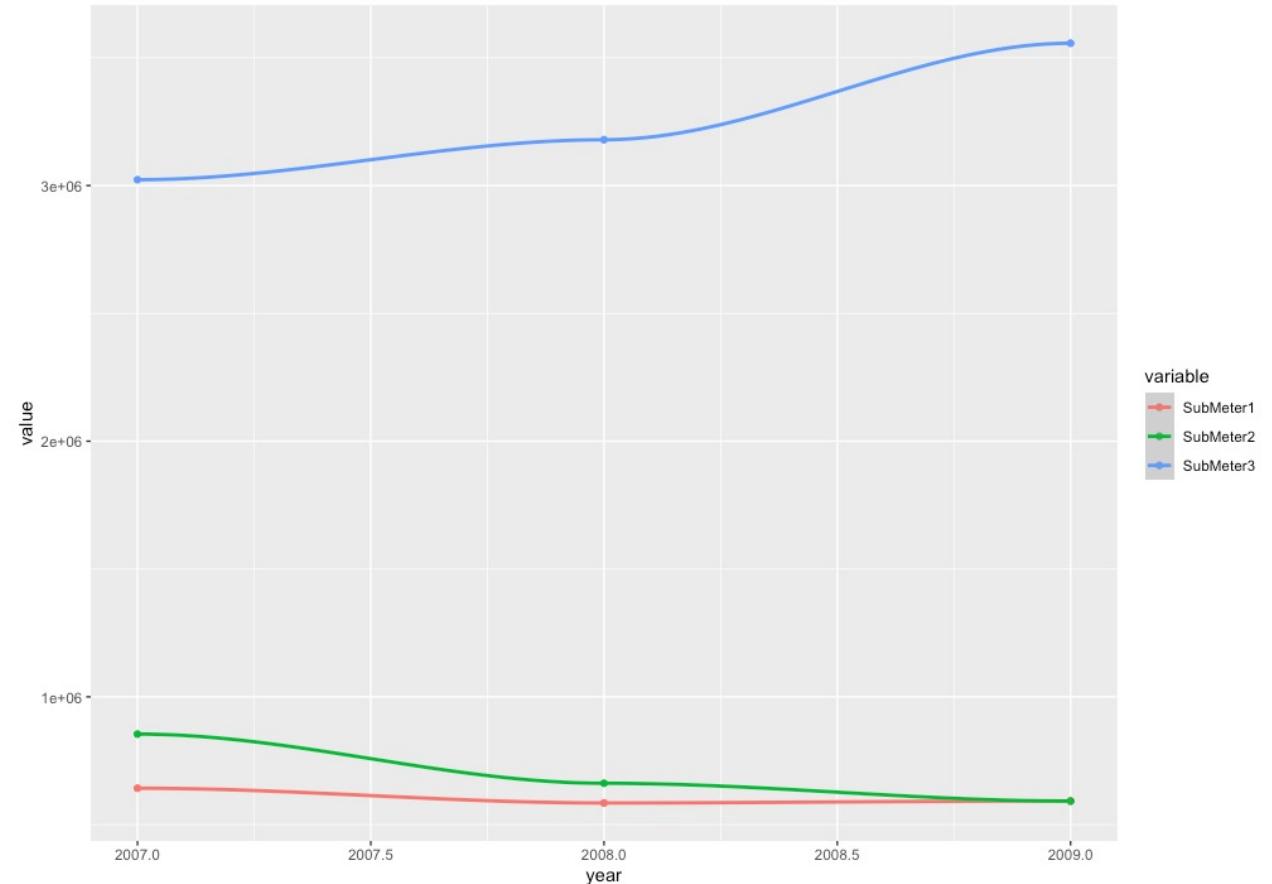
Percentage of total power use over an entire year by each sub-meter.

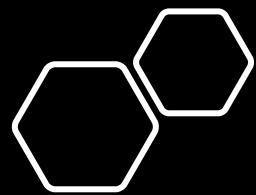
- About 70% of the energy comes from submeter 3
- Submeter 1 and 2 are used 30% each.



# Submeters aggregated and grouped by year

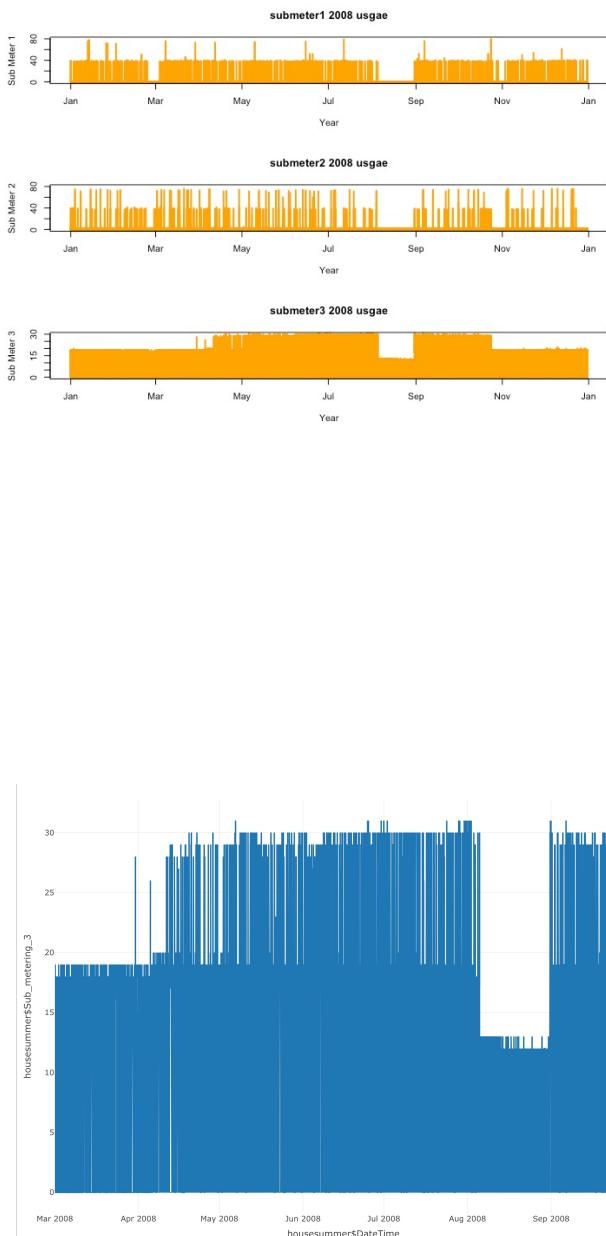
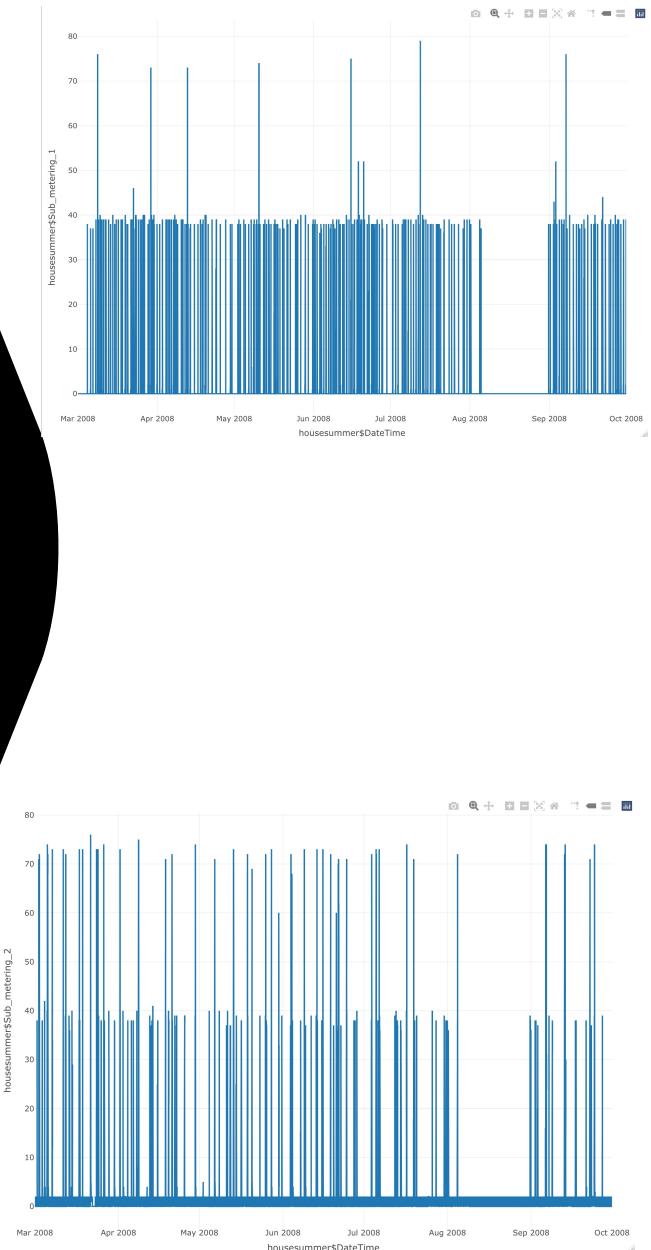
- Submeter 3 is consuming more power over the years.
- Submeter 1 is mostly constant over the years.
- Submeter 2 use declined over the years.



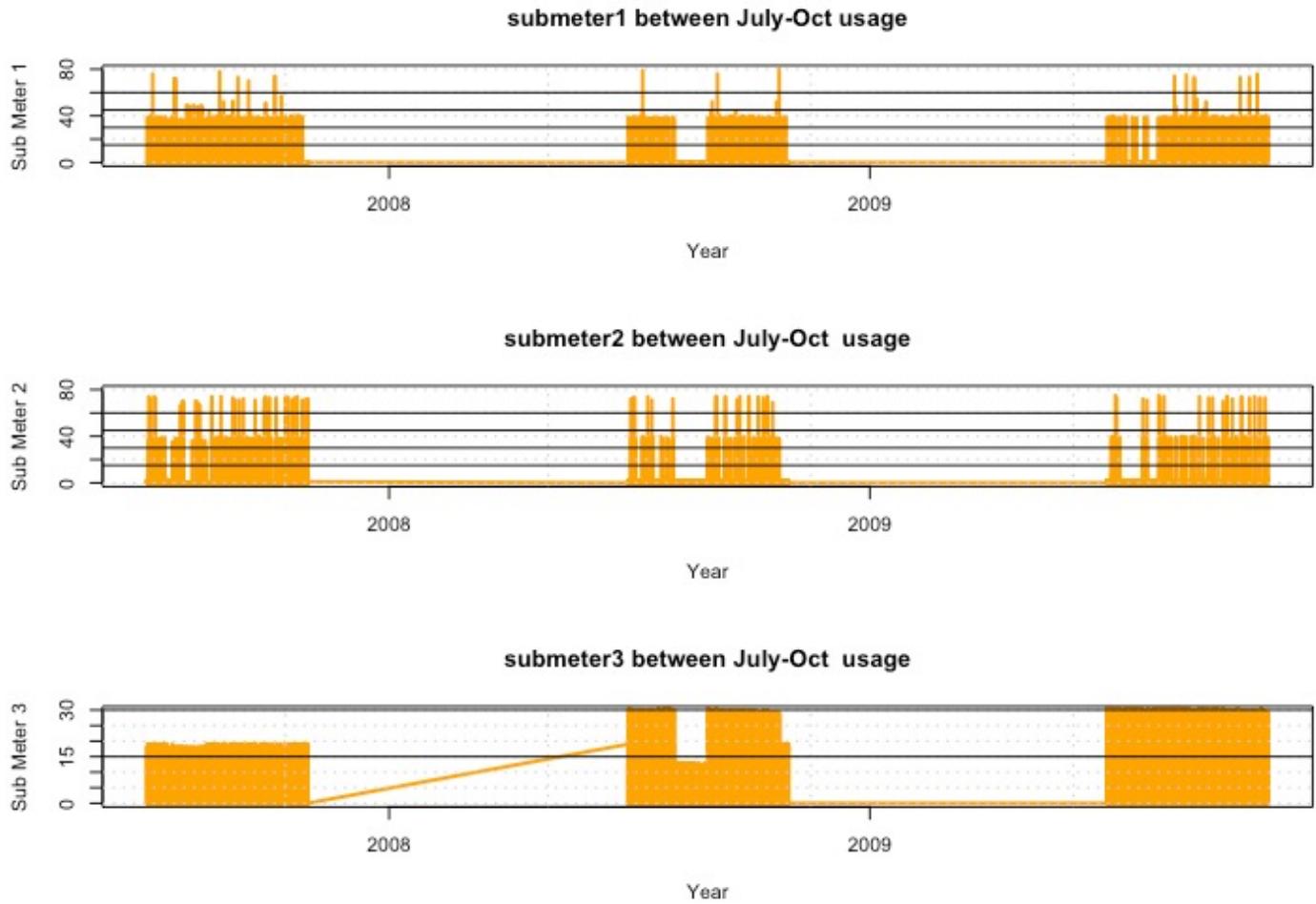


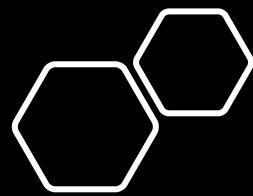
## Pattern in Summer of 2008 – month between March to September

- We can see a clear gap between August and September of 2008.
- Submeter 3 is below 10 of its normal usage.



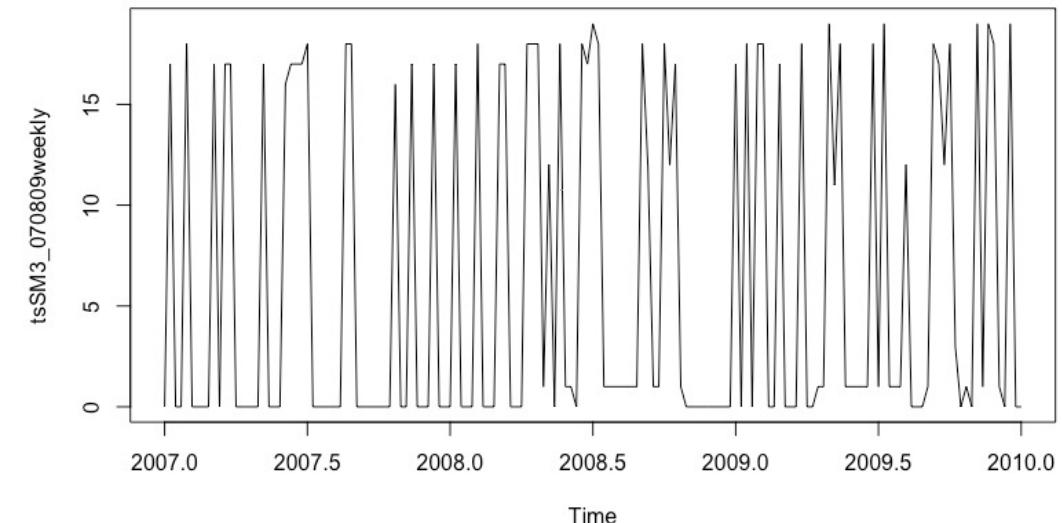
After observing the trend over years between July to October , we can clearly state that August to September 2008, the submeters were not used as much as other years



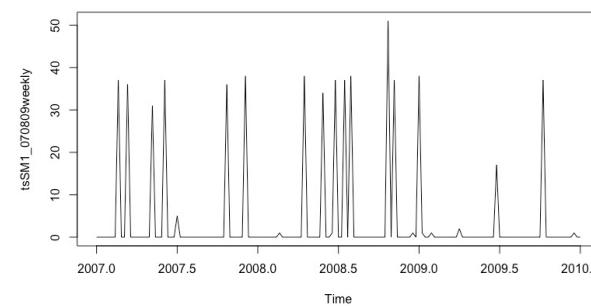


# Time Series- Weekly

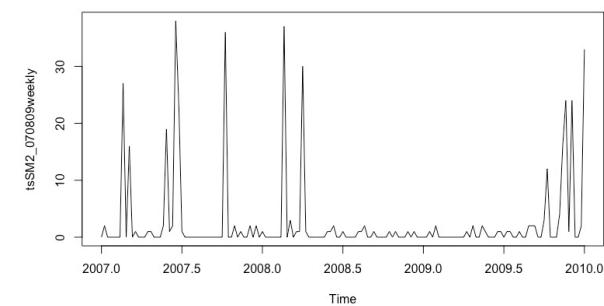
- With a 52-frequency observation in a year, we see the granularity of the submeters data.



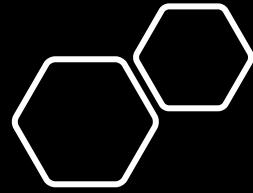
Sub 3



Sub 1

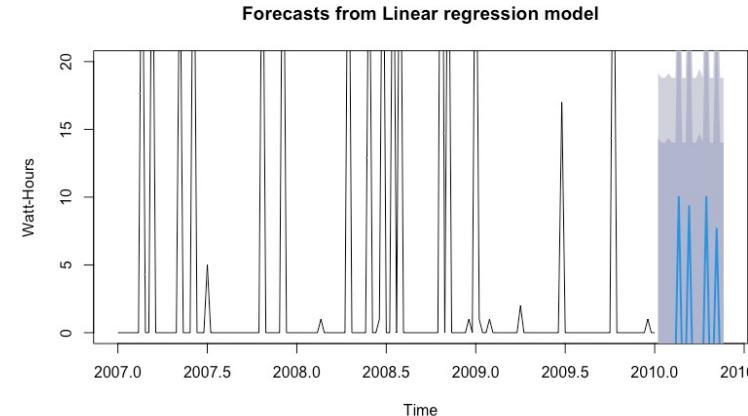


Sub 2



# Forecast

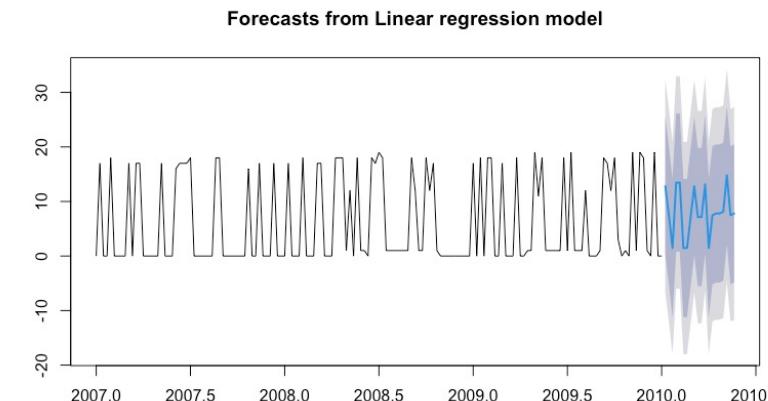
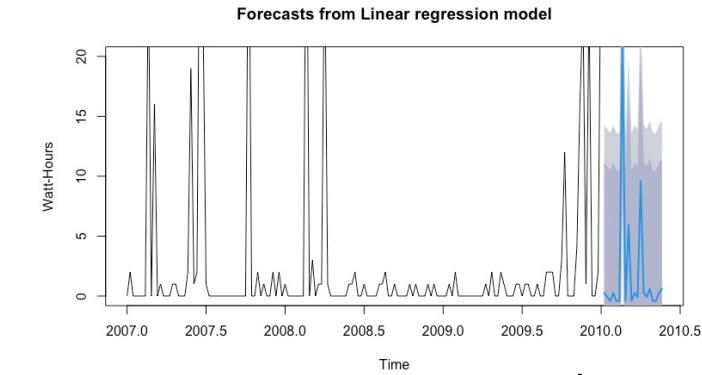
- With `tslm()` and `forecast` functions we forecast the 2010 data and see much heavier pattern in all sub meters with low r squared and RMSE value.

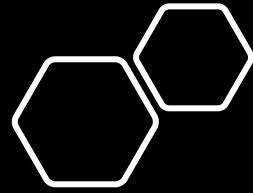


df\_summary x

Filter

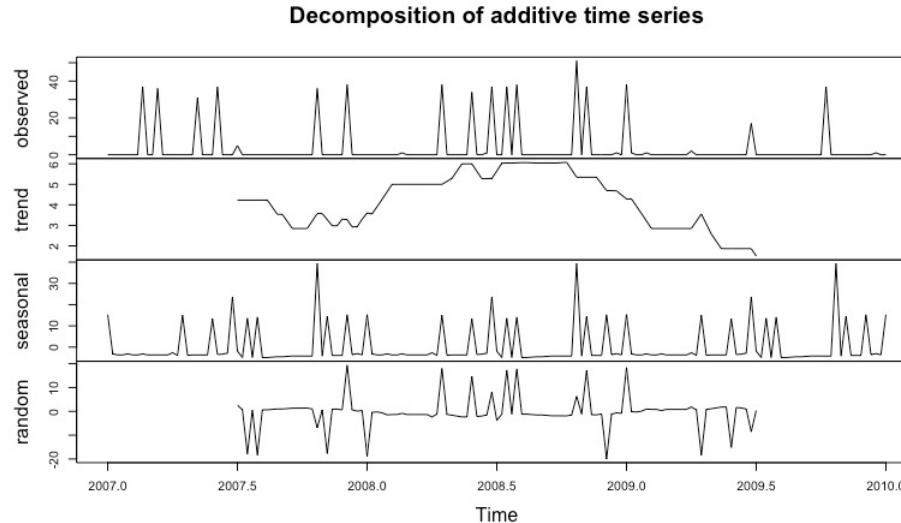
Meter	RSquared	RMSE
1 Submeter 1	0.346	8.98
2 Submeter 2	0.371	5.86
3 Submeter 3	0.278	6.82



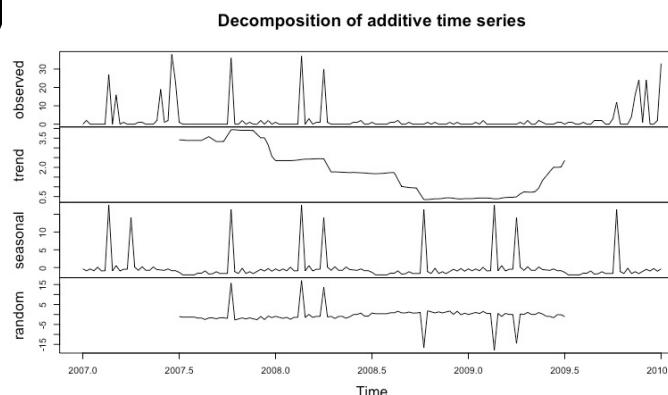


# Decomposing a Seasonal Time Series

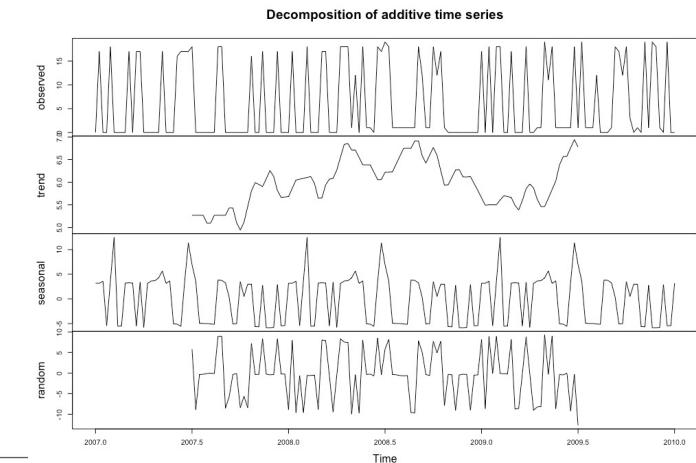
- Decomposing the trend, seasonal, and random component we can view the data and interpret useful information.
- We see Sub 3 , has an upward and downward trend usage.
- Sub 1 also has upward and downward trend with lot variations.
- Sub 2 seems different; it was upward more than downward and picks up.



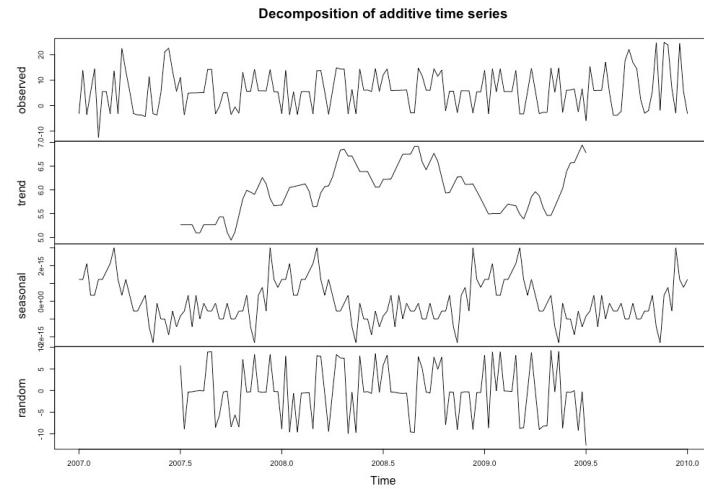
Sub 3



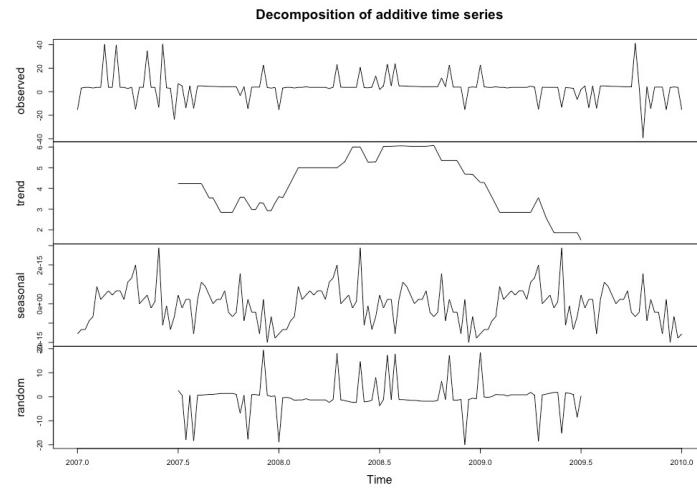
Sub 2



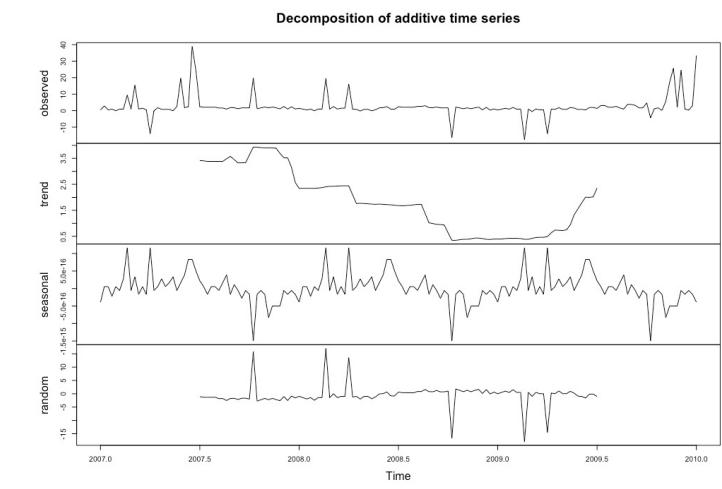
Sub 1



Sub 3



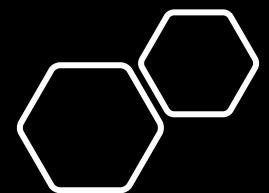
Sub 1



Sub 2

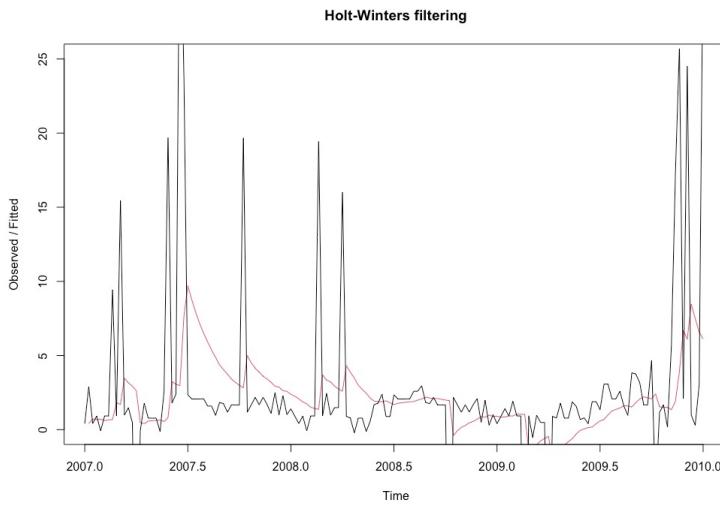
# Seasonal Adjusting

- Removing Season component, we still see the same trend in all submeters.

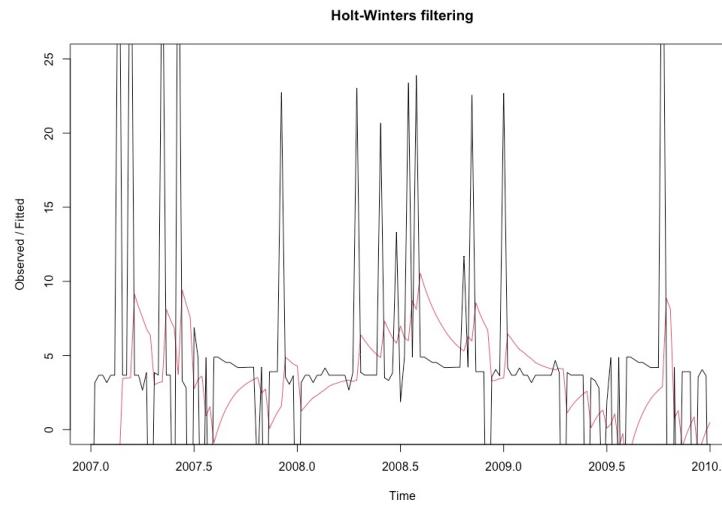


# Holt Winters Exponential Smoothing & Plot

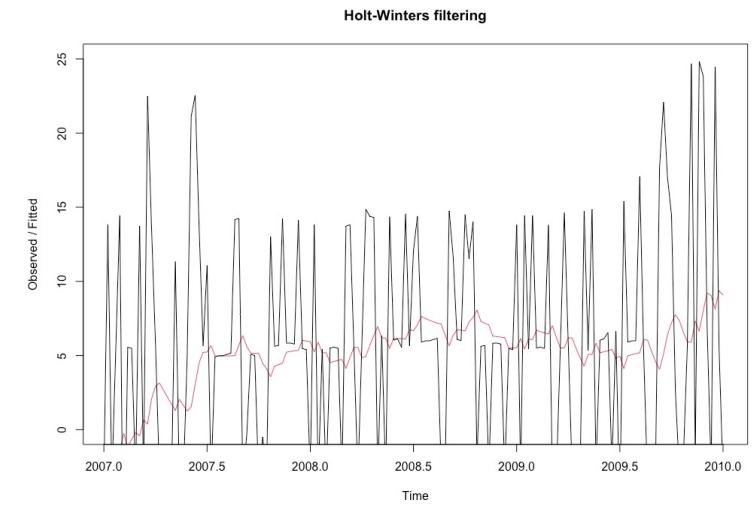
- The holt smoothing for each submeter clearly shows the forecast in red and most of it is away from the real observations confirming that the energy use of this house is higher than what it needs it to be.



SUB 2



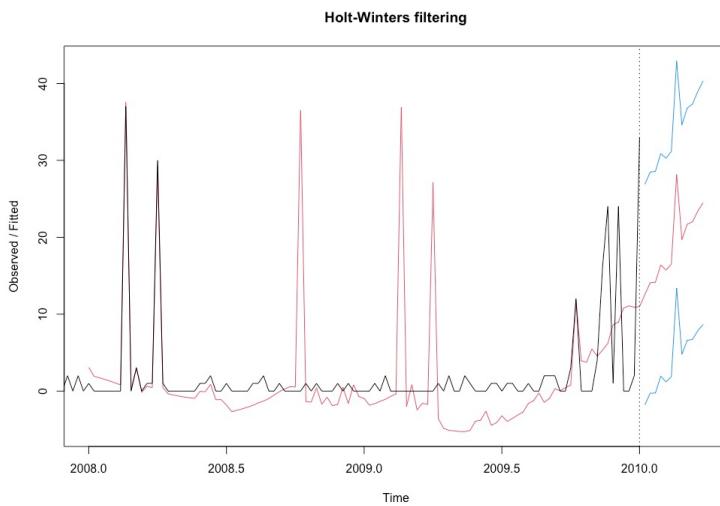
SUB 1



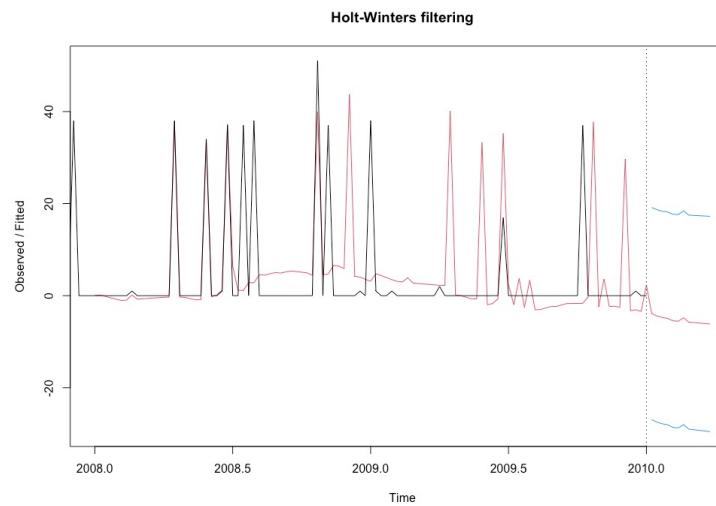
SUB 3

## Holt Winter and plot the fitted data

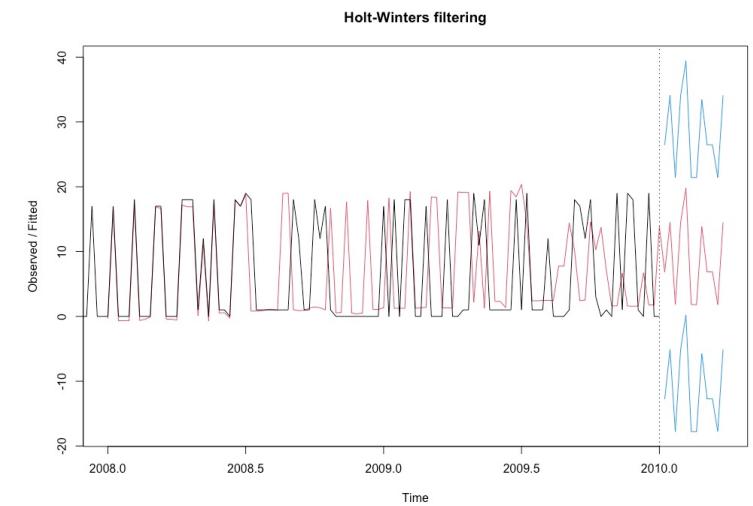
- we calculate the forecast for 12 months with a confidence interval of .95 and plot the forecast together with the actual and fitted values.



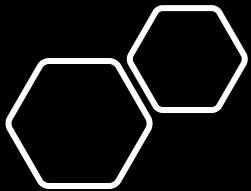
SUB 2



SUB 1



SUB 3



## Conclusions/Recommendations

- Based on Slide 13 and 14 we can clearly conclude that the house was not occupied between **August** and **September** of 2008.
- Based on the slide 16, the submeter 1 should consume less than 10 KWH in a given period of time. And Submeter 2 and 3 should be less than 20 KWH.
- From slide 17 and 18, the house does not utilize the seasonality and keeps consuming power in their own terms, if they are educated on energy use that would help change this pattern.
- This house appliances in each submeter must be inspected for energy star and see if changing any of their appliance will help achieve the holt winters smoothing plot in slide 19.
- According to the holt winters prediction in slide 20, we can state that an appliance in submeter 2(i.e., the laundry room)will start consuming more more power after 2010. Also, an appliance in Sub meter 1(i.e., Kitchen) may stop working completely. Submeter 3 will keep working but will be over consume power compared to the trend predicted.