Capstone Project

appliances energy prediction

BY [NARAYAN SINGH PARMAR
SUMIT GAIKWAD
SAGAR JAIN]

Introduction

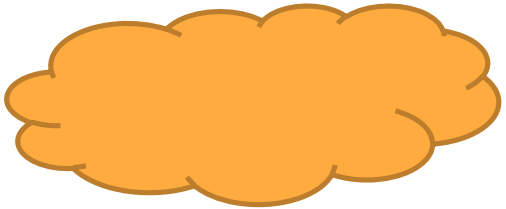
In the era of smart homes, the ability to predict energy consumption not only save money for users but also help in generating money for the user by giving excess energy back to Grid (in case of solar panels usage). In this case regression analysis will be used to predict Appliance energy usage based on data collected from various sensors. The energy prediction will come under supervised machine learning task aiming to Appliance energy consumption for a house based on factors like temperature, humidity & pressure. Many techniques, Gradient descent algorithm, and linear regression (in built function) have been applied to credit predict the energy consumption.

Dataset

The dataset (Appliances Energy Prediction)

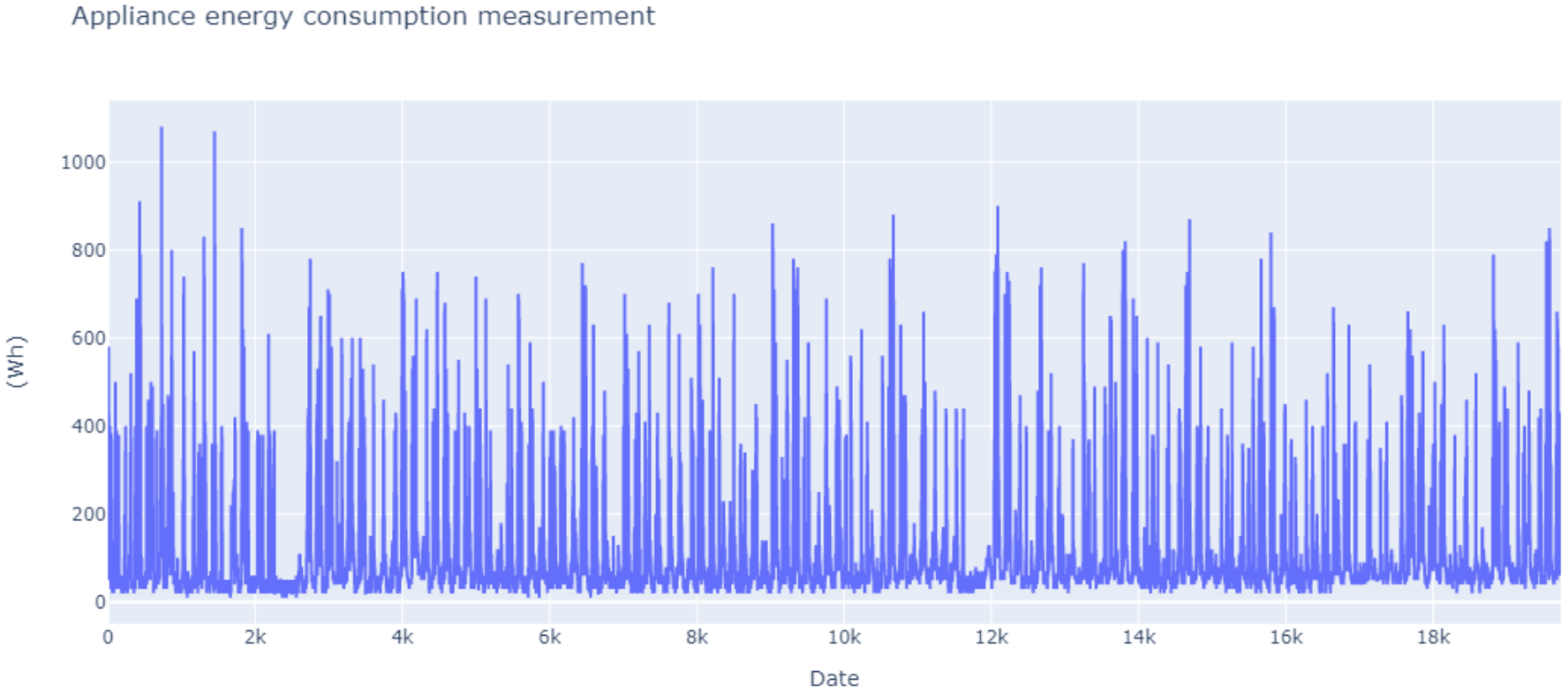
The data set is at 10 min for about 4.5 months. The house temperature and humidity conditions were monitored with a ZigBee wireless sensor network. Each wireless node transmitted the temperature and humidity conditions around 3.3 min. The dataset has 19735 rows and 29 variables include lights, date, Temperature and Humidity in various places in the house, pressure etc., The number of missing values and null values is zero. Number of weekdays is 14263 and weekend (Saturday and Sunday) is 5472.

Observations

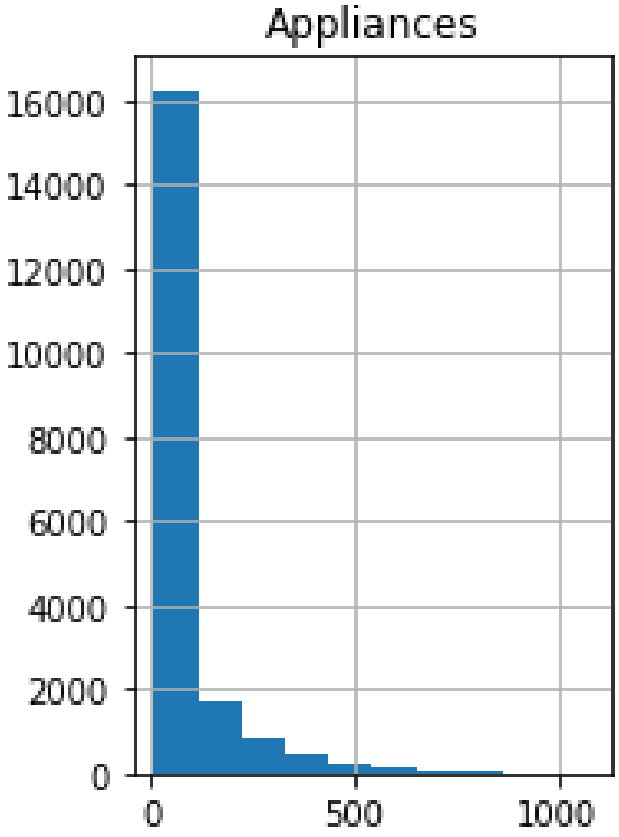


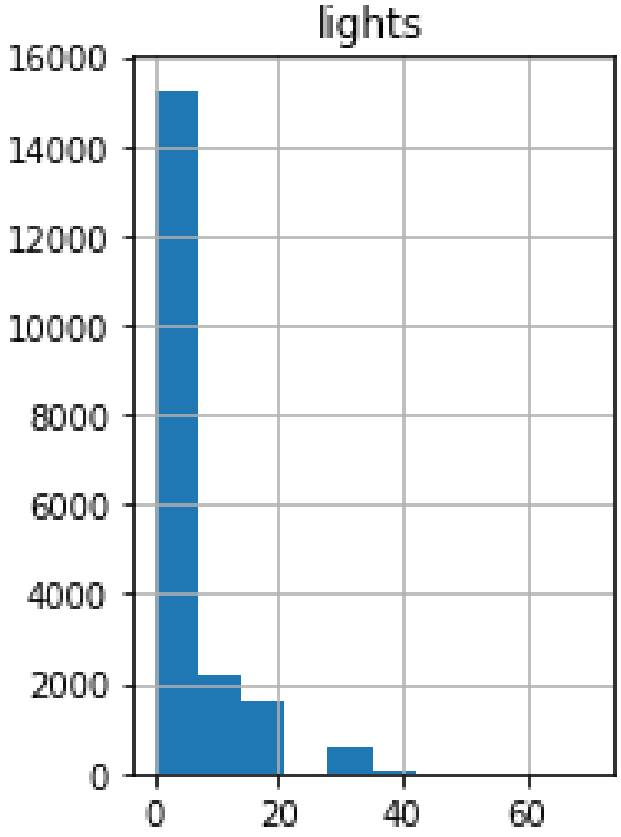
- Temperature columns - Temperature inside the house varies between 14.89 Deg & 29.85 Deg , temperatire outside (T6) varies between -6.06 Deg to 28.29 Deg . The reason for this variation is sensors are kept outside the house
- Humidiy columns - Humidity inside house varies is between 20.60% to 63.36% with exception of RH_5 (Bathroom) and RH_6 (Outside house) which varies between 29.82% to 96.32% and 1% to 99.9% respectively.
- Appliances - 75% of Appliance consumption is less than 100 Wh . With the maximum consumption of 1080 Wh , there will be outliers in this column and there are small number of cases where consumption is very high
- Lights column - Intially I believed lights column will be able to give useful information . With 11438 0 (zero) enteries in 14801 rows , this column will not add any value to the model . I believed light consumption along with humidity level in a room will give idea about human presence in the room and hence its impact on Appliance consumption. Hence for now , I will dropping this column

Energy Consumption:

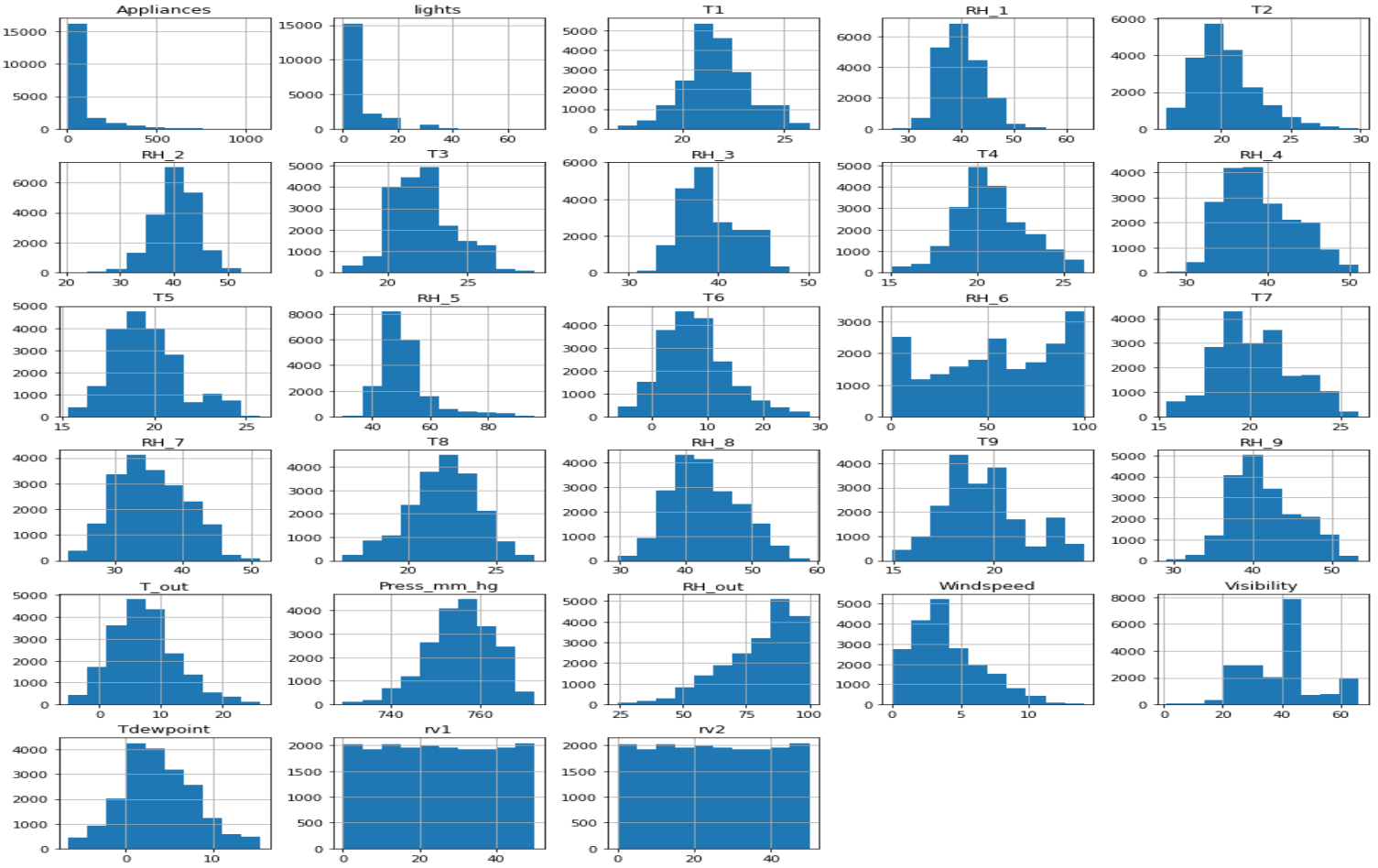


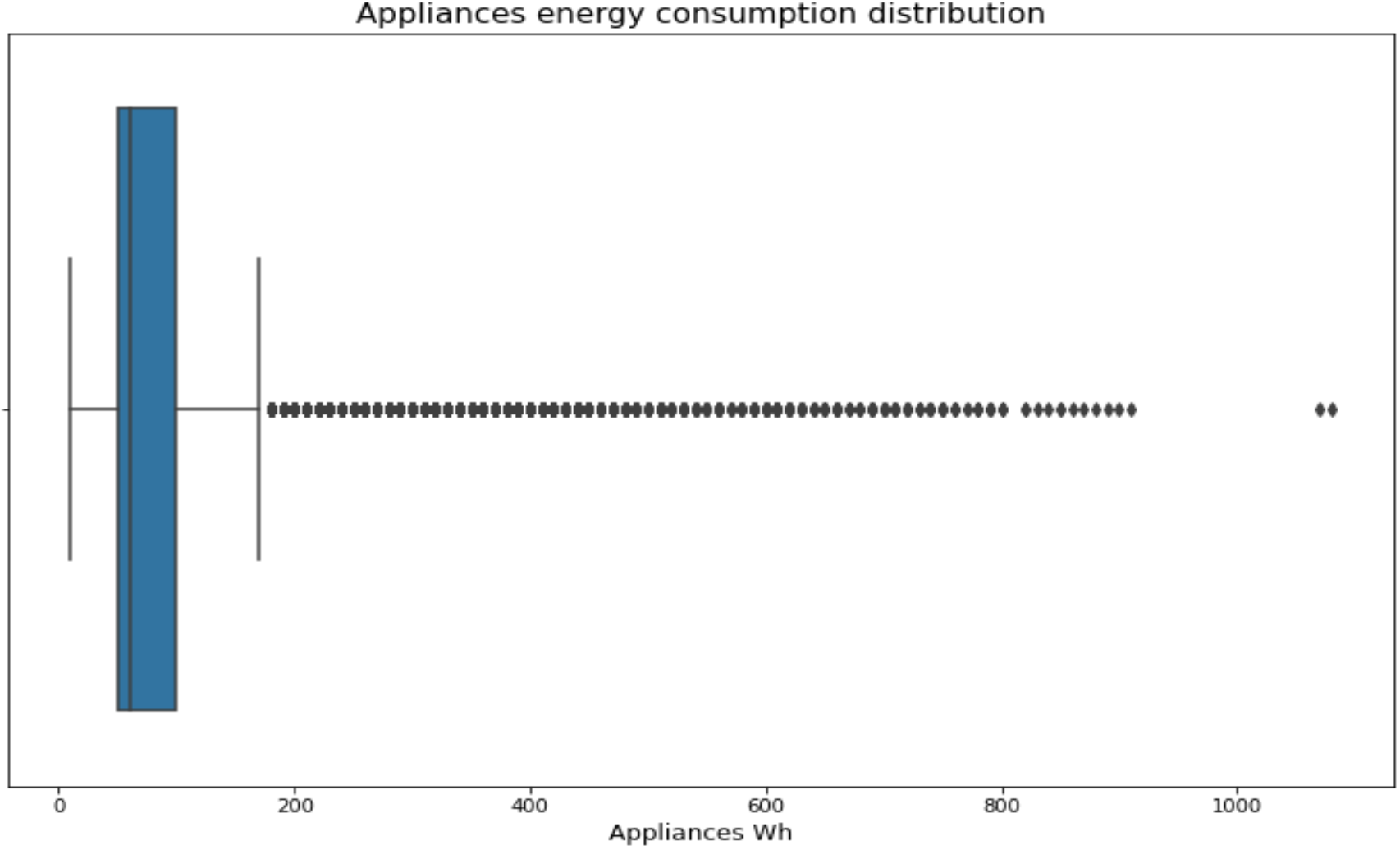
Appliance = lights

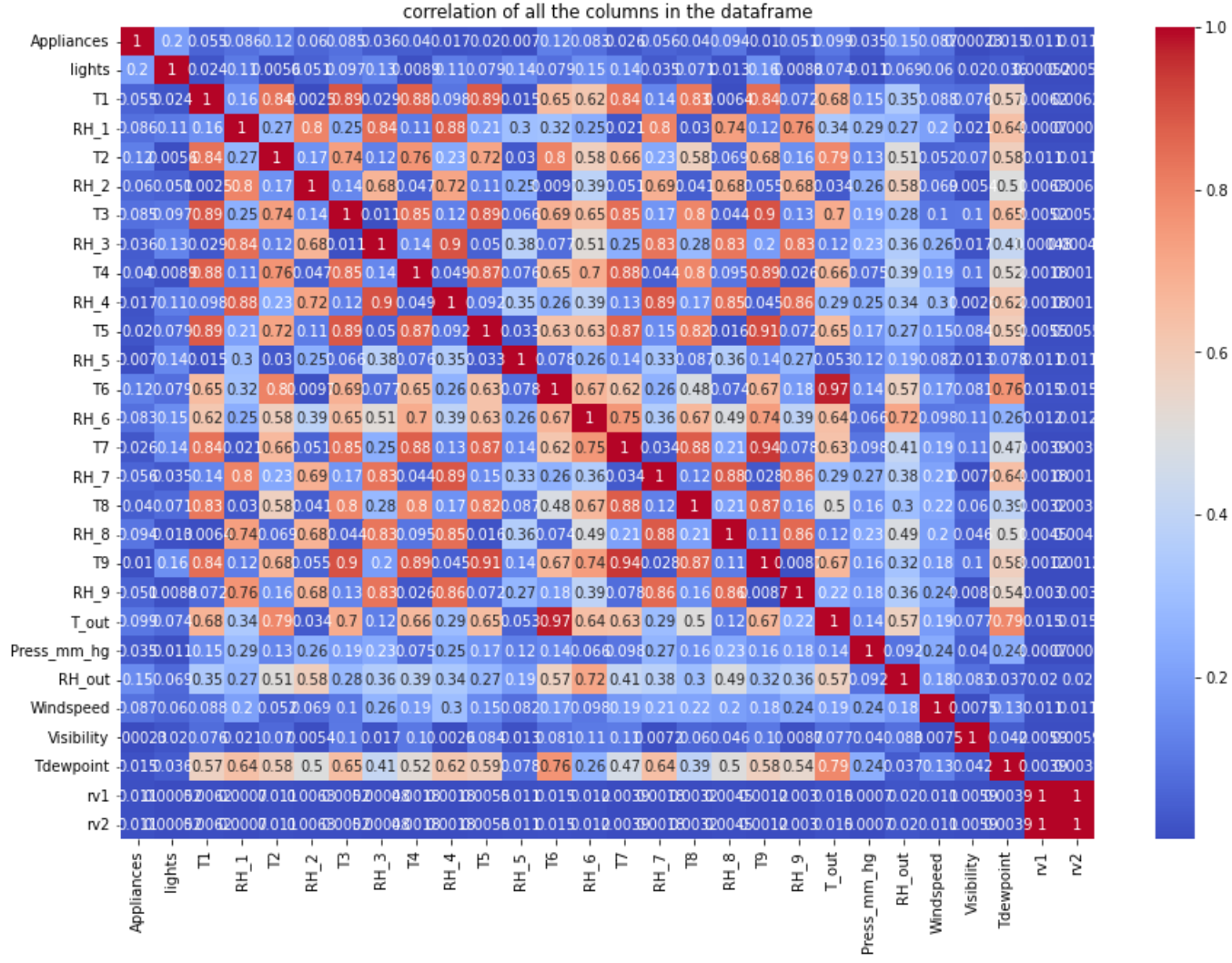




Feature distribution




Appliances energy consumption distribution
0
200
400
600
800
1000
Appliances Wh


correlation of all the columns in the dataframe
Appliances
lights
T1
RH_1
T2
RH_2
T3
RH_3
T4
RH_4
T5
RH_5
T6
RH_6
T7
RH_7
T8
RH_8
T9
RH_9
T_out
Press_mm_hg
RH_out
Windspeed
Visibility
Tdewpoint
rv1
rv2
1.0
0.8
0.6
0.4
0.2

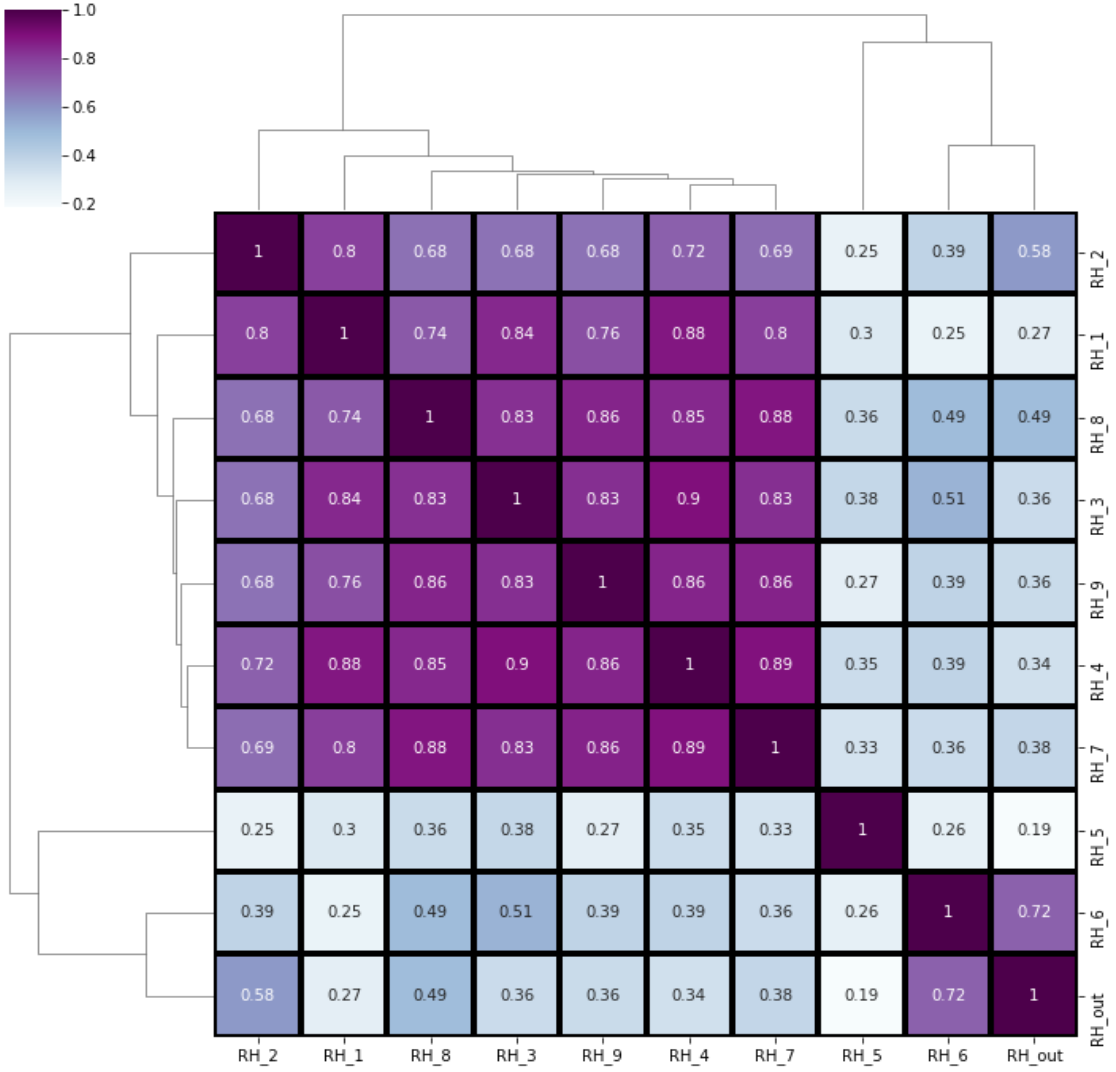
Observations based on correlation plot

All the temperature variables from T1-T9 and T_out have positive correlation with the target Appliances . For the indoortemperatures, the correlations are high as expected, since the ventilation is driven by the HRV unit and minimizes air tempera-ture differences between rooms. Four columns have a high degree of correlation with T9 - T3,T5,T7,T8 also T6 & T_Out has high correlation (both temperatures from outside) . Hence T6 & T9 can be removed from training set as information provided by them can be provided by other fields.

Weather attributes - Visibility, Tdewpoint, Press_mm_hg have low correlation values

Humidity - There are no significantly high correlation cases (> 0.9) for humidity sensors.

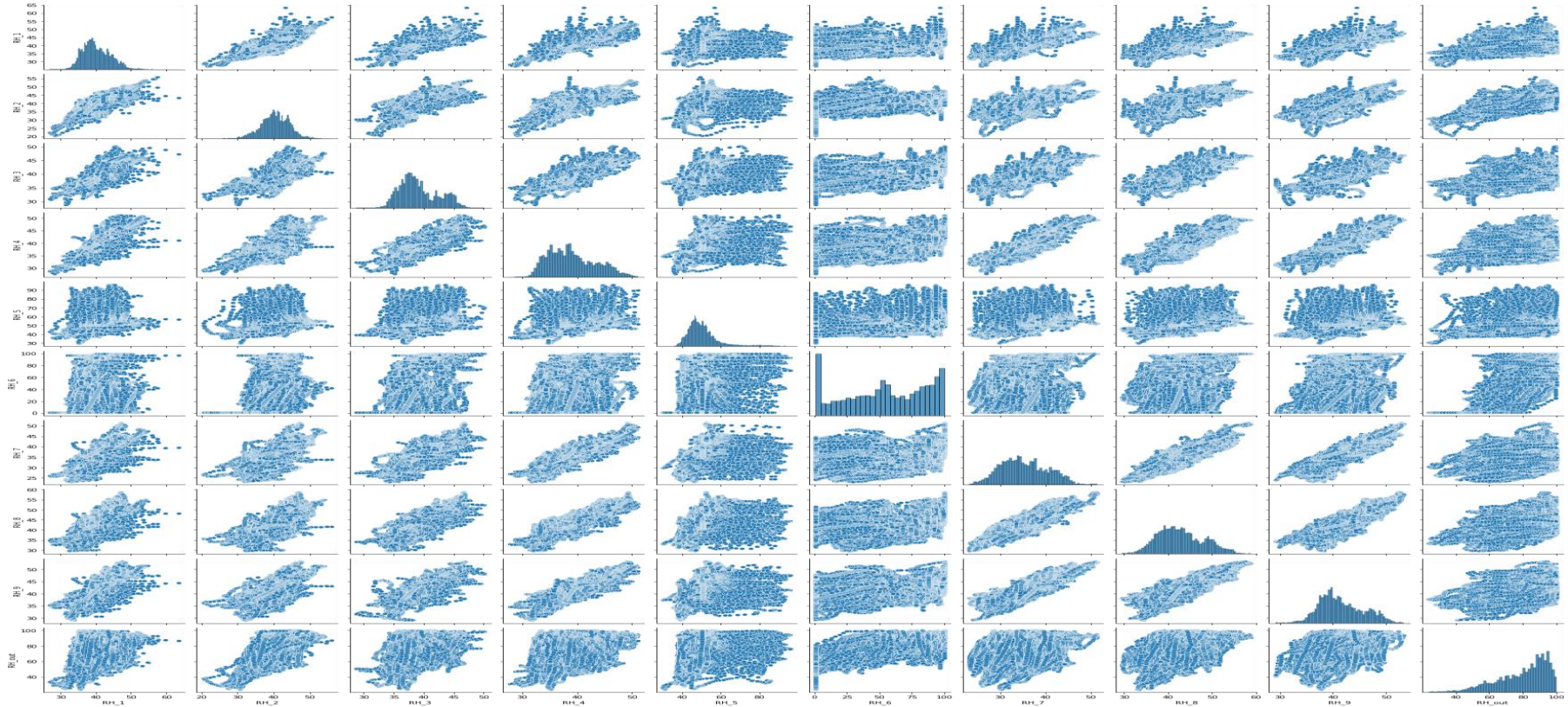
Random variables have no role to play



Correlation matrix for humdity indoor and outdoor

RH_1, Humidity in kitchen area, in % RH_2, Humidity in living room area, in % RH_3, Humidity in laundry room area, in % RH_4, Humidity in office room, in % RH_5, Humidity in bathroom, in % RH_6, Humidity outside the building (north side), in % RH_7, Humidity in ironing room, in % RH_8, Humidity in teenager room 2, in % RH_9, Humidity in parents room, in % RH_out, Humidity outside (from Chievres weather station), in %

Pariplot for humidity



Conclusion:

This analysis shows that how much impact weather conditions have in our day-today energy consumption in home. Whenever the Atomospheric pressure is low or high, there's a evident change in the way enegry is consumed