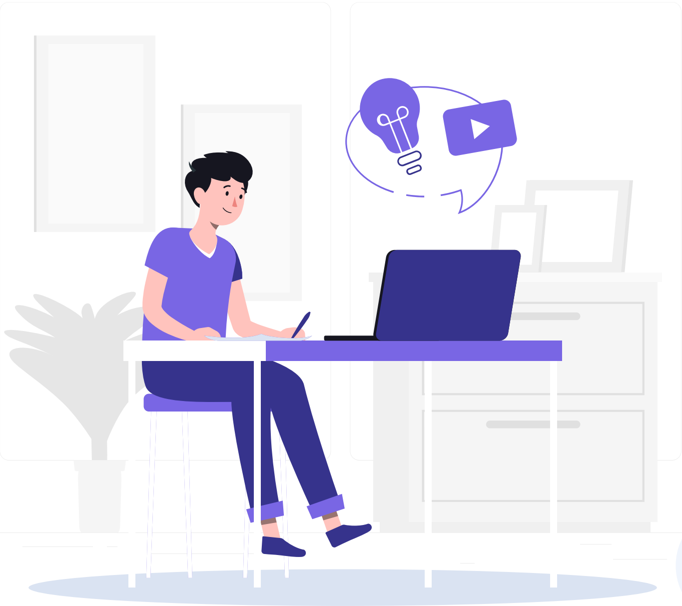
Occupancy Detection

Created and Presented
By
Cristina Sahoo



Applications

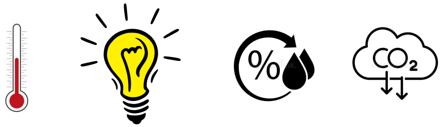
- Recent studies and measurements *[12-15]* report energy savings of 30% to 42% with accurate occupancy determination

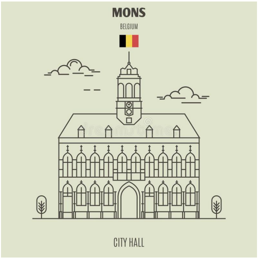
- When occupancy data was used as an input for HVAC control algorithms, energy savings were as high as 80% *[16]*

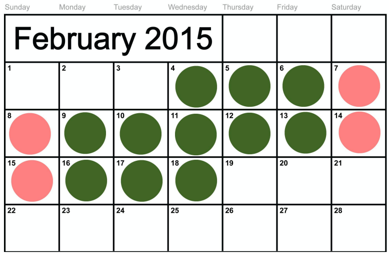
- A system that could accurately detect the presence of the occupants without using a camera is very interesting due to privacy concerns

- Other applications for occupancy detection include security and determination of building occupant behaviors

Problem Statement



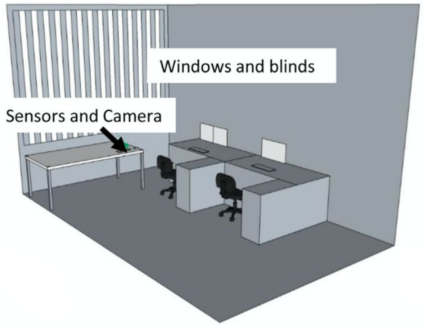


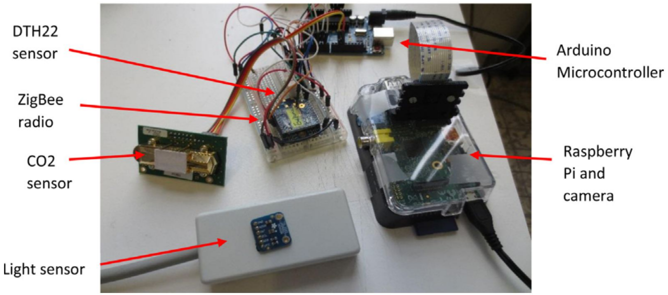


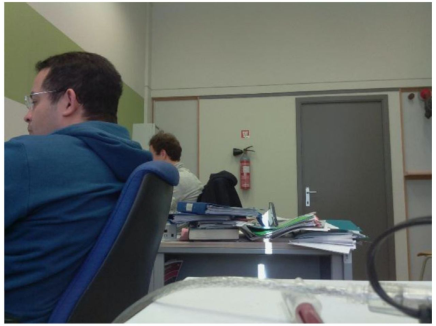
- Identify trends over time and correlation between environment variables and occupancy
- Identify algorithms/models with best accuracy score for predicting occupancy

Data Source and Collection

Occupancy Detection Dataset - UCI Machine Learning Repository



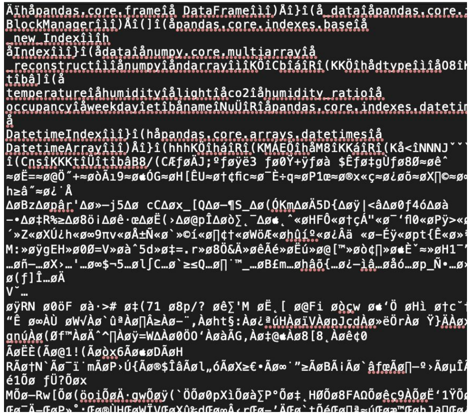




Data Cleaning and Preprocessing

```
"date","Temperature","Humidity","Light","CO2","HumidityRatio","Occupancy"
"1","2015-02-04 17:51:00",23.18,27.272,426,721.25,0.00479298817650529,1
"2","2015-02-04 17:51:59",23.15,27.2675,429.5,714,0.00478344094931065,1
"3","2015-02-04 17:53:00",23.15,27.245,426,713.5,0.00477946352442199,1
"4","2015-02-04 17:54:00",23.15,27.2,426,708.25,0.00477150882608175,1
"5","2015-02-04 17:55:00",23.1,27.2,426,704.5,0.00475699293331518,1
"6","2015-02-04 17:55:59",23.1,27.2,419,701,0.00475699293331518,1
"7","2015-02-04 17:57:00",23.1,27.2,419,701.666666666667,0.00475699293331518,1
"8","2015-02-04 17:57:59",23.1,27.2,419,699,0.00475699293331518,1
"9","2015-02-04 17:58:59",23.1,27.2,419,689.333333333333,0.00475699293331518,1
"10","2015-02-04 18:00:00",23.075,27.175,419,688,0.00474535071966655,1
"11","2015-02-04 18:01:00",23.075,27.15,419,690.25,0.00474095189694268,1
"12","2015-02-04 18:02:00",23.1,27.1,419,691,0.00473937073052061,1
"13","2015-02-04 18:03:00",23.1,27.1666666666667,419,683.5,0.00475111875560951,1
"14","2015-02-04 18:04:00",23.05,27.15,419,687.5,0.00473373179708025,1
"15","2015-02-04 18:04:59",23,27.125,419,686,0.00471494214590473,1
"16","2015-02-04 18:06:00",23,27.125,418.5,680.5,0.00471494214590473,1
"17","2015-02-04 18:07:00",23,27.2,0,681.5,0.0047280779496687,0
"18","2015-02-04 18:08:00",22.945,27.29,0,685,0.00472795137178073,0
"19","2015-02-04 18:08:59",22.945,27.39,0,685,0.0047454083970941,0
"20","2015-02-04 18:10:00",22.89,27.39,0,689,0.00472950615591001,0
"21","2015-02-04 18:10:59",22.89,27.39,0,689.5,0.00472950615591001,0
"22","2015-02-04 18:11:59",22.89,27.39,0,689,0.00472950615591001,0
"23","2015-02-04 18:13:00",22.89,27.445,0,691,0.00473907551474663,0
"24","2015-02-04 18:14:00",22.89,27.5,0,688,0.00474864516581148,0
"25","2015-02-04 18:15:00",22.89,27.5,0,689.5,0.00474864516581148,0
"26","2015-02-04 18:16:00",22.79,27.445,0,689,0.00471022404680625,0
"27","2015-02-04 18:16:59",22.79,27.5,0,685.666666666667,0.00471973499914952,0
"28","2015-02-04 18:17:59",22.79,27.5,0,687,0.00471973499914952,0
```

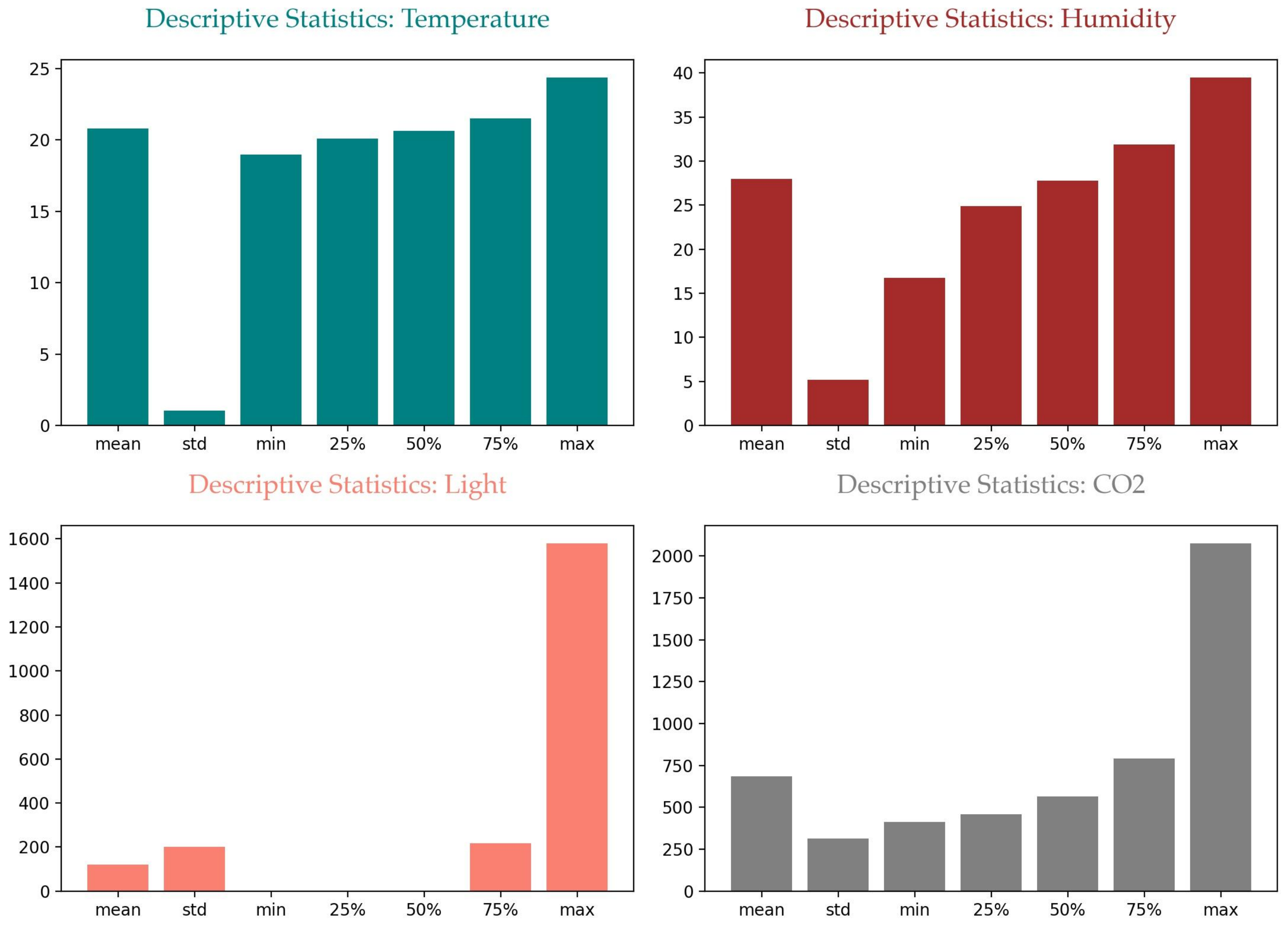
date	temperature	humidity	light	co2	humidity_ratio	occupancy
2015-02-04 17:51:00	23.18	27.272	426.0	721.25	0.00479298817650529	1
2015-02-04 17:52:00	23.15	27.2675	429.5	714.0	0.00478344094931065	1
2015-02-04 17:53:00	23.15	27.245	426.0	713.5	0.00477946352442199	1
2015-02-04 17:54:00	23.15	27.2	426.0	708.25	0.00477150882608175	1
2015-02-04 17:55:00	23.1	27.2	426.0	704.5	0.004756992933315180	1
2015-02-04 17:56:00	23.1	27.2	419.0	701.0	0.004756992933315180	1
2015-02-04 17:57:00	23.1	27.2	419.0	701.666666666667	0.004756992933315180	1
2015-02-04 17:58:00	23.1	27.2	419.0	699.0	0.004756992933315180	1
2015-02-04 17:59:00	23.1	27.2	419.0	689.333333333333	0.004756992933315180	1
2015-02-04 18:00:00	23.075	27.175	419.0	688.0	0.00474535071966655	1
2015-02-04 18:01:00	23.075	27.15	419.0	690.25	0.00474095189694268	1
2015-02-04 18:02:00	23.1	27.1	419.0	691.0	0.00473937073052061	1
2015-02-04 18:03:00	23.1	27.1666666666667	419.0	683.5	0.00475111875560951	1
2015-02-04 18:04:00	23.05	27.15	419.0	687.5	0.0047337317970825	1
2015-02-04 18:05:00	23.0	27.125	419.0	686.0	0.00471494214590473	1
2015-02-04 18:06:00	23.0	27.125	418.5	680.5	0.00471494214590473	1
2015-02-04 18:07:00	23.0	27.2	0.0	681.5	0.004728077949668770	0
2015-02-04 18:08:00	22.945	27.29	0.0	685.0	0.00472795137178073	0
2015-02-04 18:09:00	22.945	27.39	0.0	685.0	0.0047454083970941	0
2015-02-04 18:10:00	22.89	27.39	0.0	689.0	0.00472950615591001	0
2015-02-04 18:11:00	22.89	27.39	0.0	689.5	0.00472950615591001	0
2015-02-04 18:12:00	22.89	27.39	0.0	689.0	0.00472950615591001	0
2015-02-04 18:13:00	22.89	27.445	0.0	691.0	0.00473907551474663	0
2015-02-04 18:14:00	22.89	27.5	0.0	688.0	0.00474864516581148	0
2015-02-04 18:15:00	22.89	27.5	0.0	689.5	0.00474864516581148	0
2015-02-04 18:16:00	22.79	27.445	0.0	689.0	0.00471022404680625	0



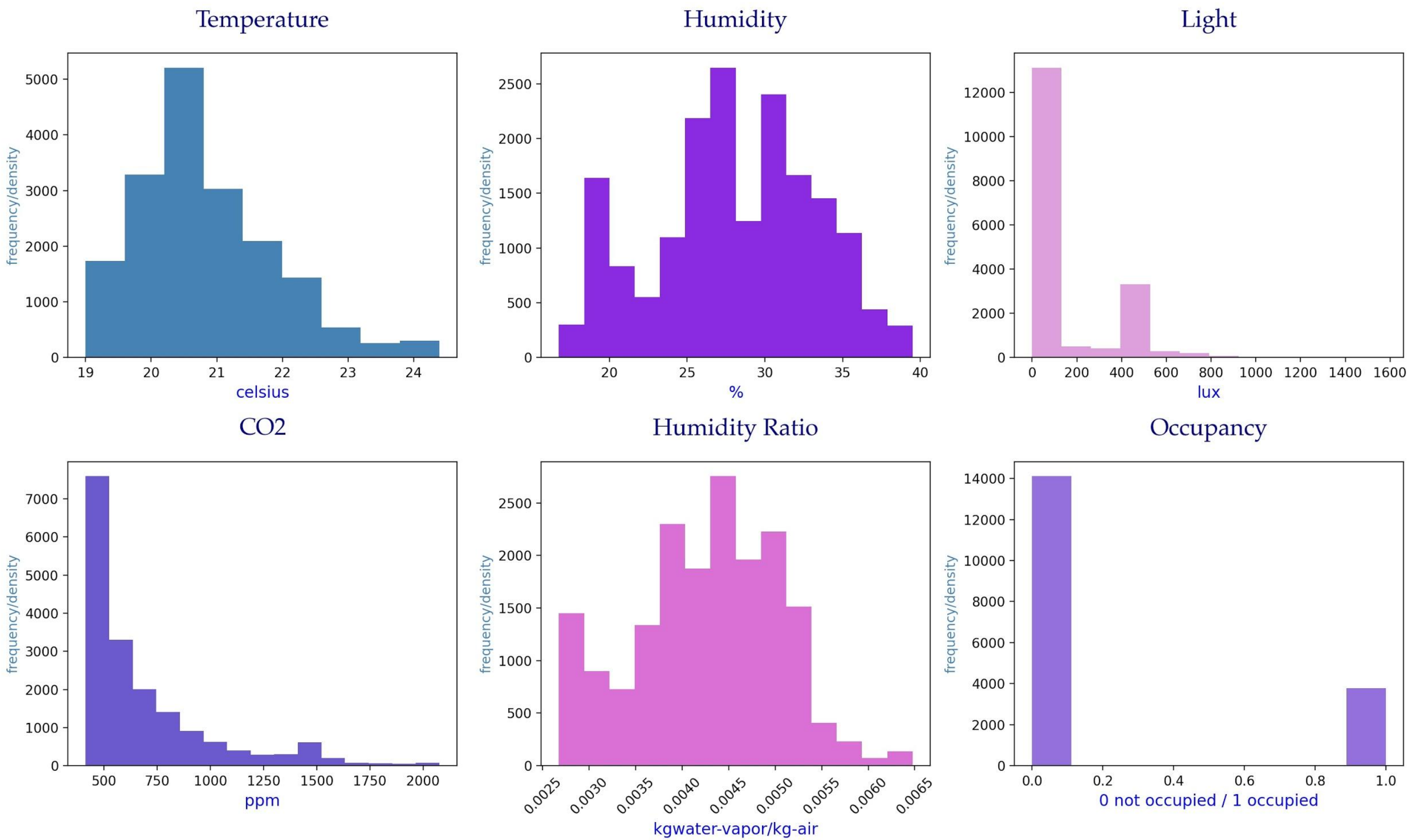
- Remove or impute null values
- Remove/impute outliers
- Remove duplicate observations
- Rename columns
- Round dates to nearest minute
- Sort by date, chronologically
- Index by date to allow time series type exploration of the data
- Create the “weekday” feature

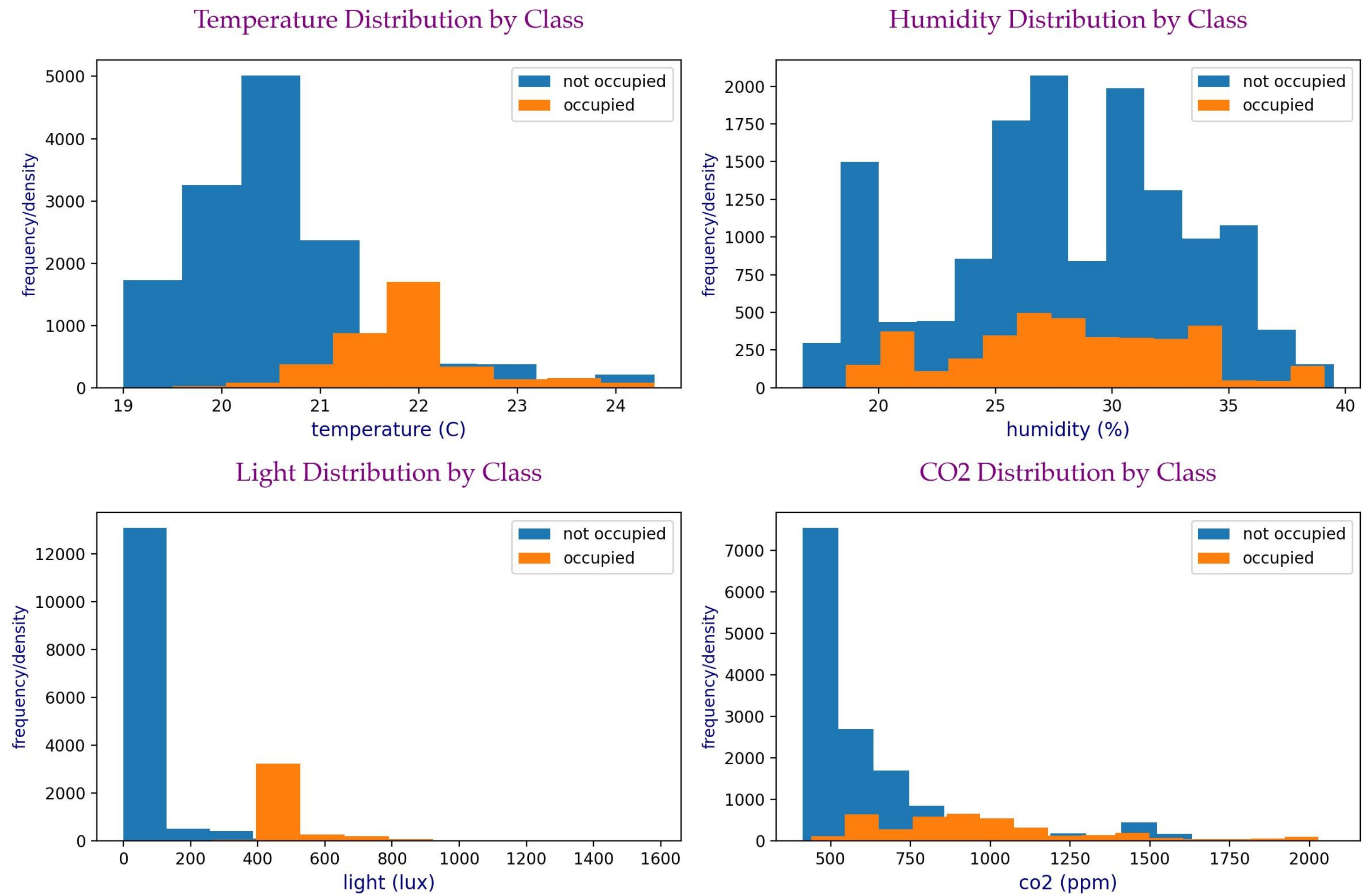
Data Dictionary

Feature Name	Feature Description	Units of Measurement or Format
date	time the observation was recorded	year-month-day hour:minute:second
temperature	temperature recorded	Celsius
humidity	relative humidity recorded	%
light	light recorded at time of observation	Lux
co2	CO2 measured at the time of observation	ppm, parts per million
humidity_ratio	derived quantity from temperature and relative humidity	kgwater-vapor/kg-air
occupancy	status of room occupancy	0 for not occupied, 1 for occupied status
weekday	indicates if the timestamp is weekday or weekend	1 for weekday, 0 for weekend


Descriptive Statistics: Temperature
0
5
10
15
20
25
mean
std
min
25%
50%
75%
max
Descriptive Statistics: Humidity
0
5
10
15
20
25
30
35
40
mean
std
min
25%
50%
75%
max
Descriptive Statistics: Light
0
200
400
600
800
1000
1200
1400
1600
mean
std
min
25%
50%
75%
max
Descriptive Statistics: CO2
0
250
500
750
1000
1250
1500
1750
2000
mean
std
min
25%
50%
75%
max

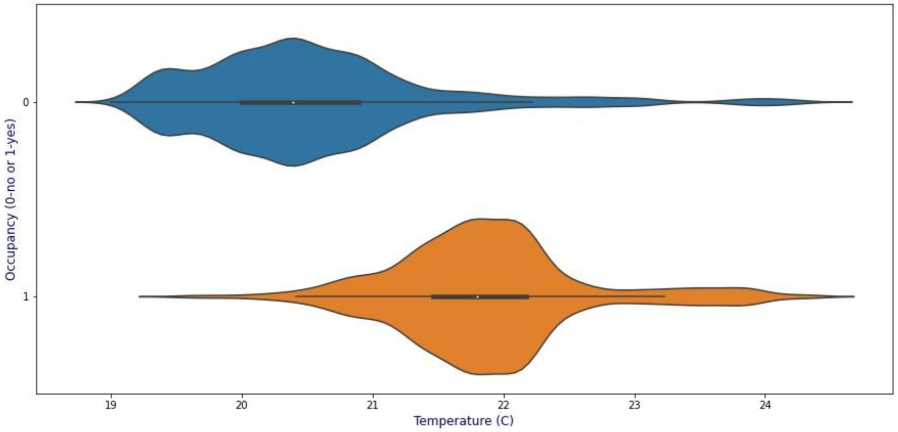
Data Distributions



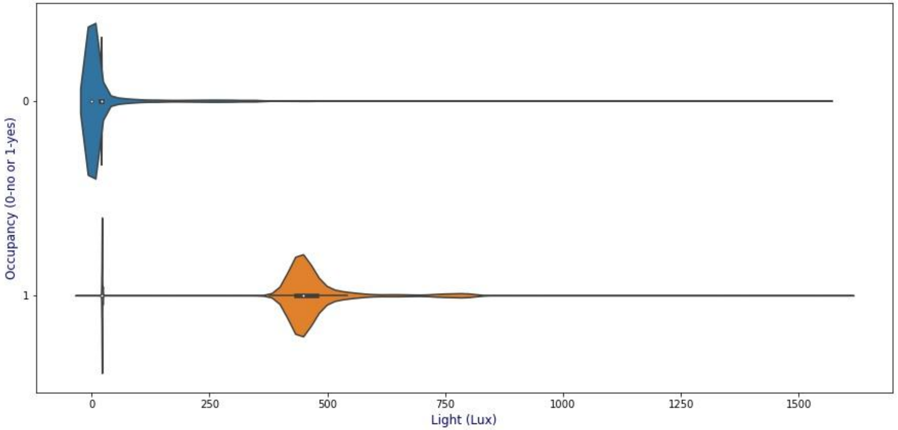


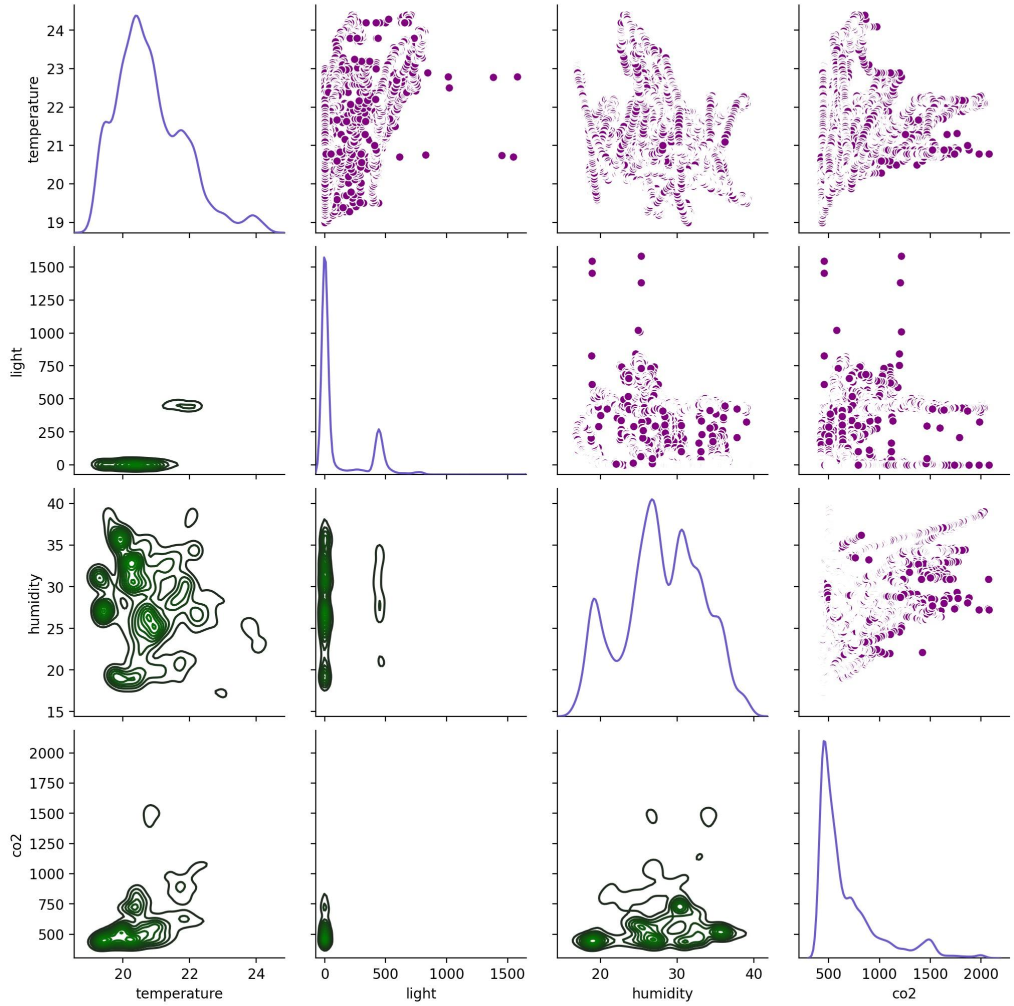
Temperature Distribution by Class
not occupied
occupied
frequency/density
temperature (C)
Humidity Distribution by Class
not occupied
occupied
frequency/density
humidity (%)
Light Distribution by Class
not occupied
occupied
frequency/density
light (lux)
CO2 Distribution by Class
not occupied
occupied
frequency/density
co2 (ppm)

Temperature Distribution by Class

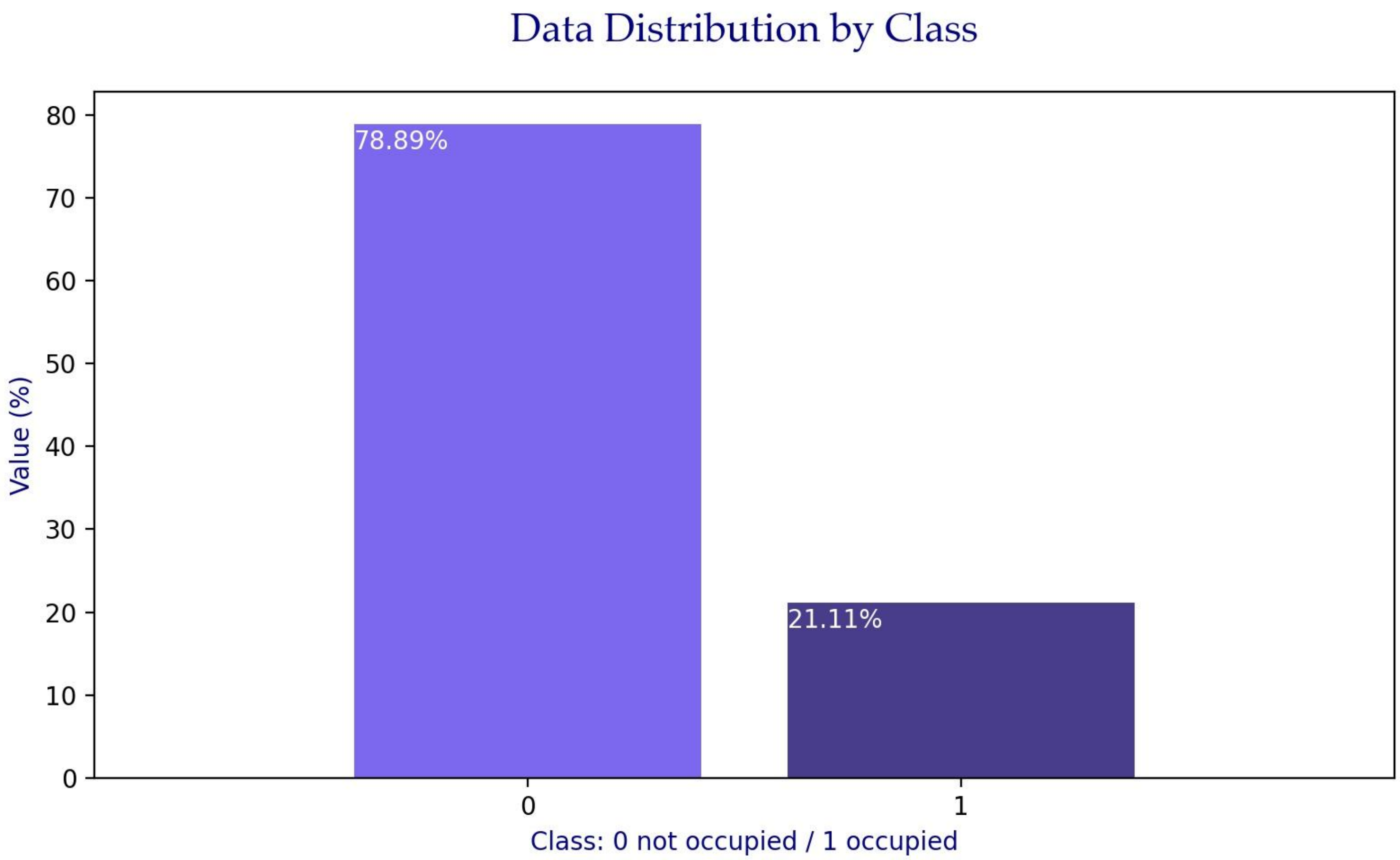


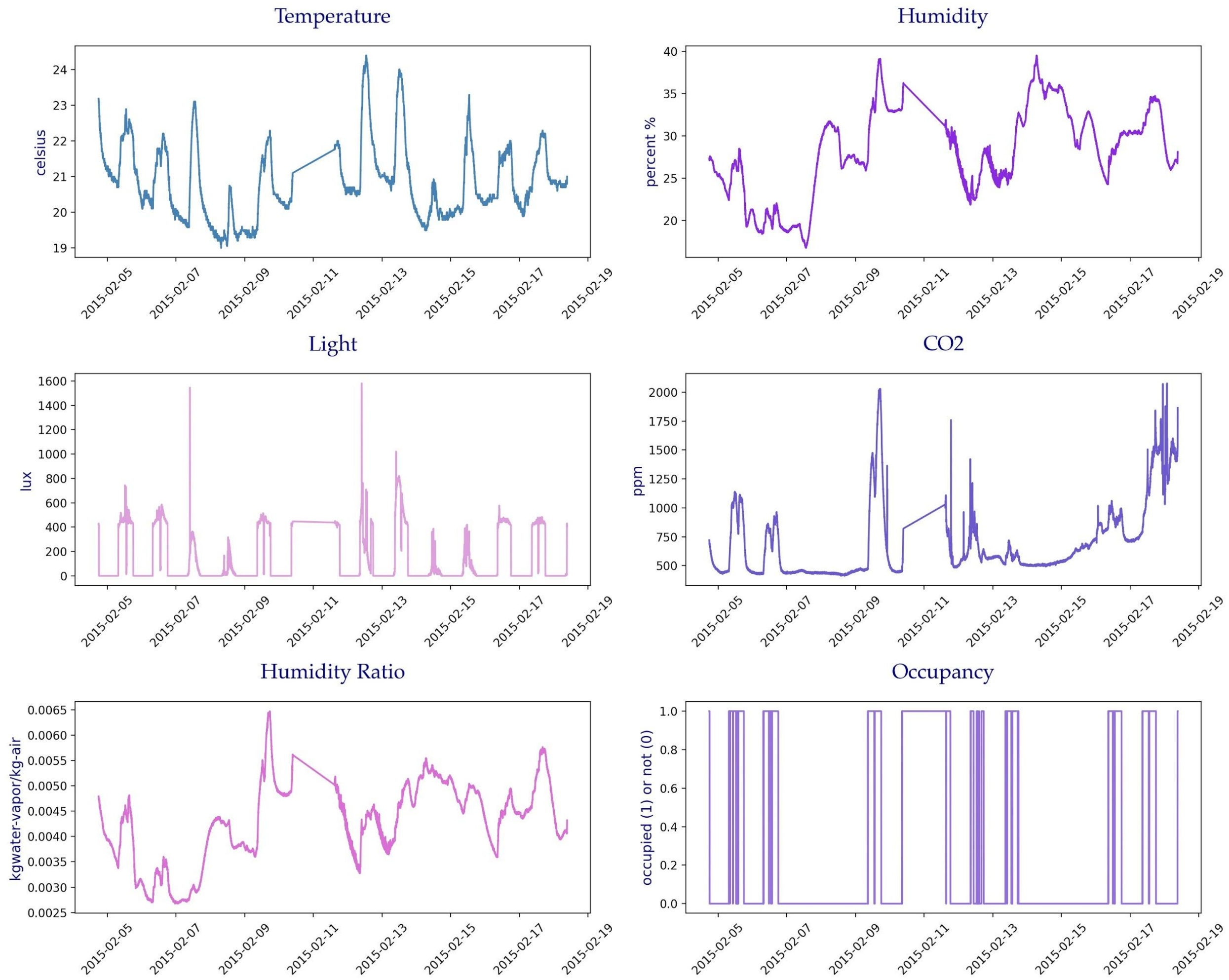
Light Distribution by Class

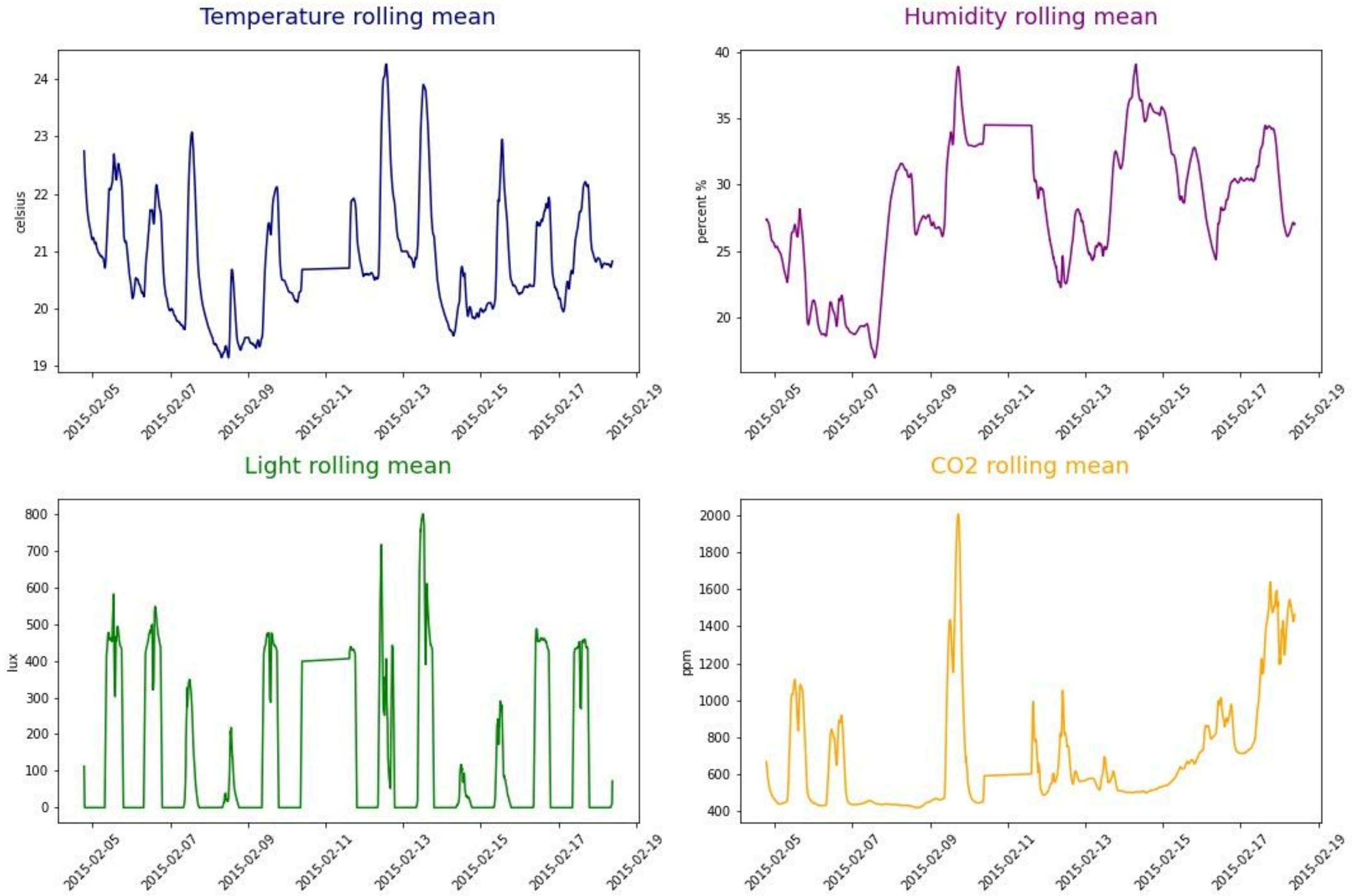


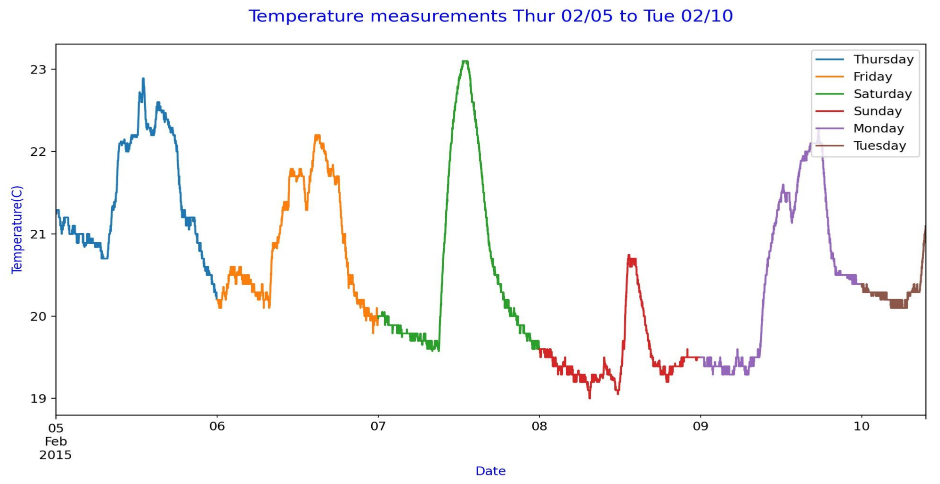


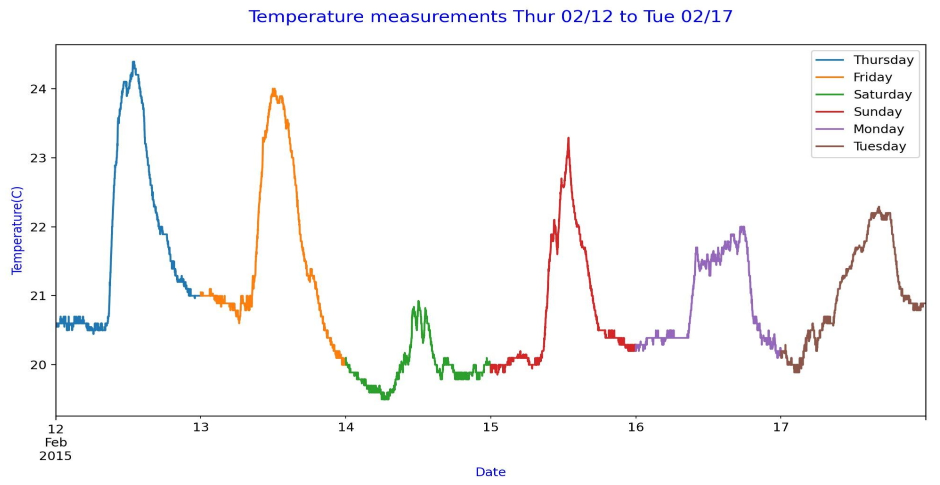
temperature
light
humidity
co2
temperature
light
humidity
co2


Data Distribution by Class
Value (%)
80
70
60
50
40
30
20
10
0
78.89%
21.11%
0
1
Class: 0 not occupied / 1 occupied

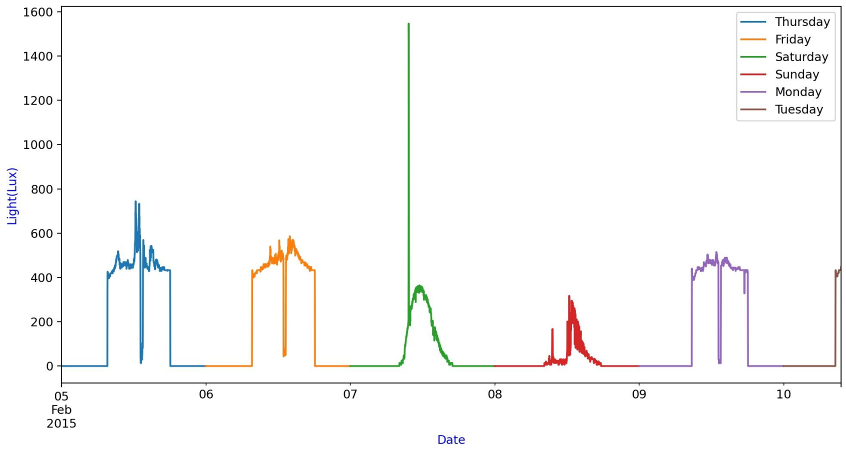

Temperature
celsius
Humidity
percent %
Light
lux
CO2
ppm
Humidity Ratio
kgwater-vapor/kg-air
Occupancy
occupied (1) or not (0)
2015-02-05
2015-02-07
2015-02-09
2015-02-11
2015-02-13
2015-02-15
2015-02-17
2015-02-19


Temperature rolling mean
celsius
Humidity rolling mean
percent %
Light rolling mean
lux
CO2 rolling mean
ppm
2015-02-05
2015-02-07
2015-02-09
2015-02-11
2015-02-13
2015-02-15
2015-02-17
2015-02-19

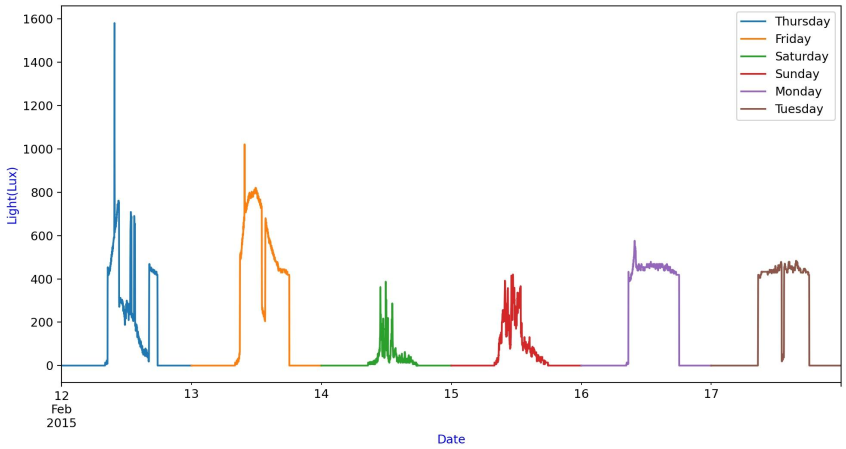

Temperature measurements Thur 02/05 to Tue 02/10
Thursday
Friday
Saturday
Sunday
Monday
Tuesday
Temperature(C)
23
22
21
20
19
05
Feb
2015
06
07
08
09
10
Date


Temperature measurements Thur 02/12 to Tue 02/17
Thursday
Friday
Saturday
Sunday
Monday
Tuesday
Temperature(C)
24
23
22
21
20
12
Feb
2015
13
14
15
16
17
Date

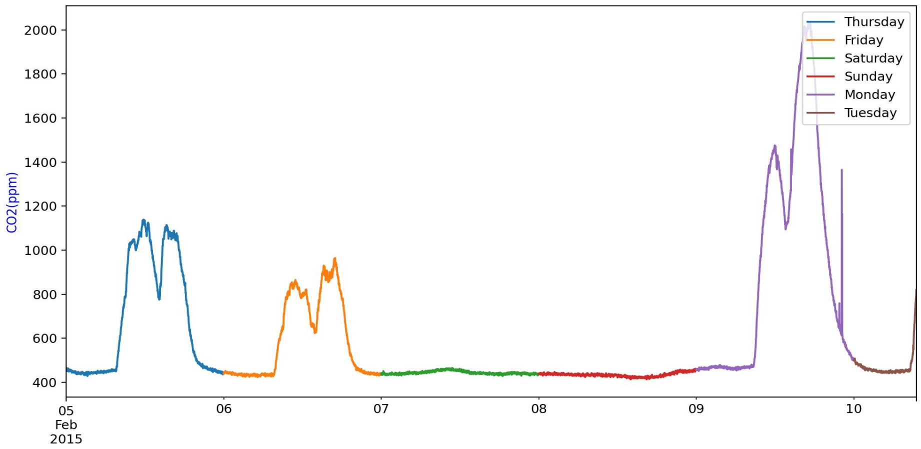
Light measurements Thur 02/05 to Tue 02/10



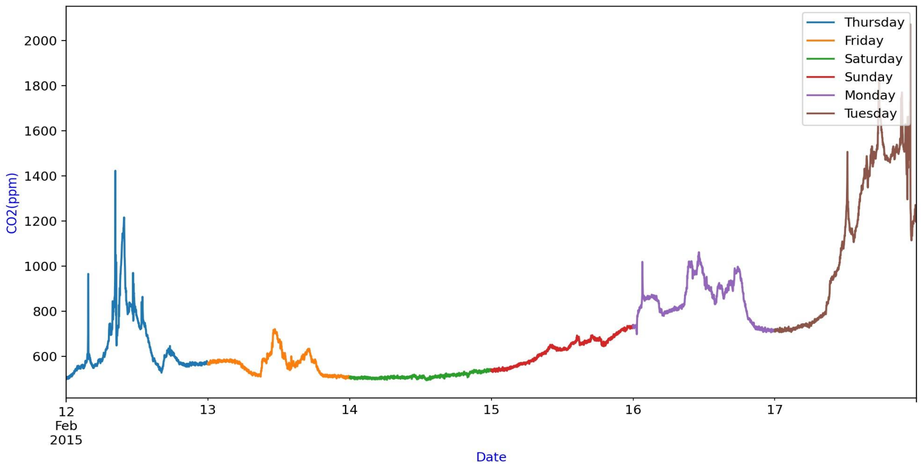
Light measurements Thur 02/12 to Tue 02/17

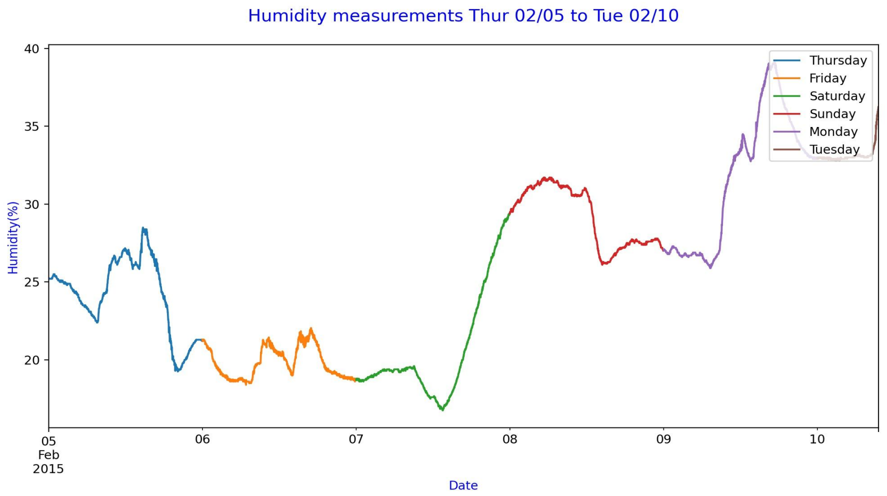


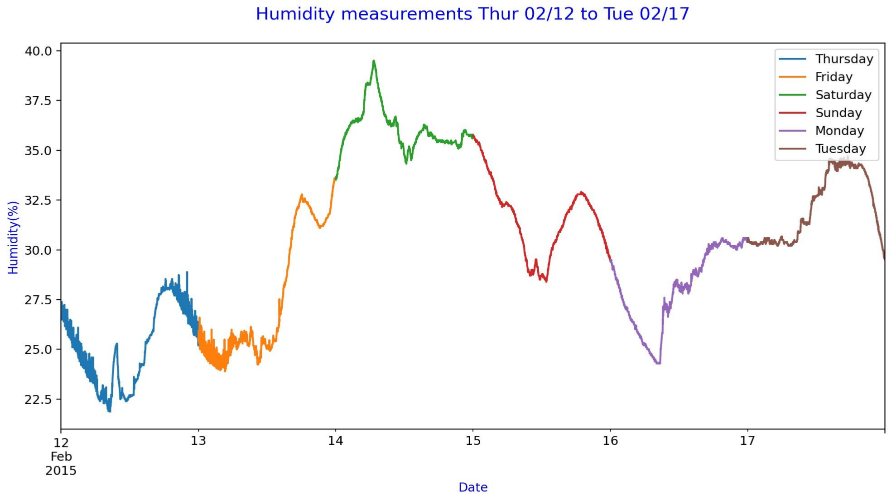
CO2 measurements Thur 02/05 to Tue 02/10



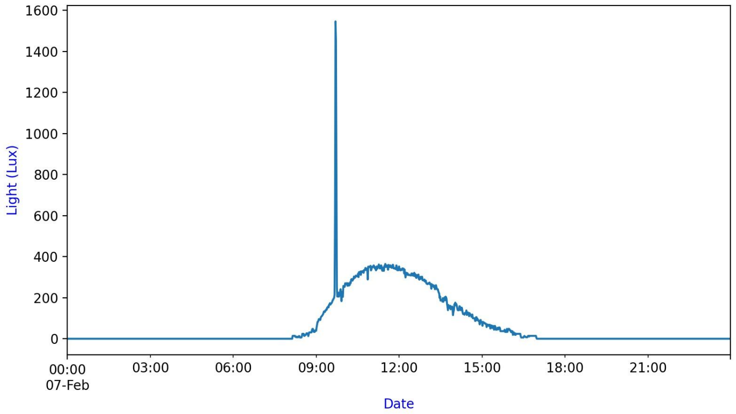
CO2 measurements Thur 02/12 to Tue 02/17



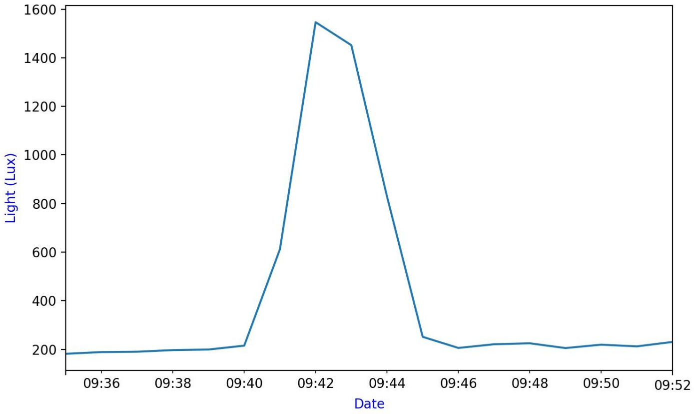

Humidity measurements Thur 02/05 to Tue 02/10
Thursday
Friday
Saturday
Sunday
Monday
Tuesday
40
35
30
25
20
Humidity(%)
05
Feb
2015
06
07
08
09
10
Date


Humidity measurements Thur 02/12 to Tue 02/17
Thursday
Friday
Saturday
Sunday
Monday
Tuesday
40.0
37.5
35.0
32.5
30.0
27.5
25.0
22.5
Humidity(%)
12
Feb
2015
13
14
15
16
17
Date

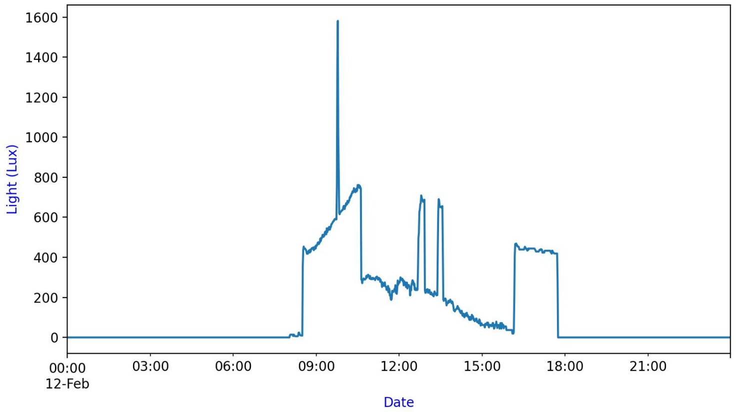
Light measurements on Saturday 2015-02-07



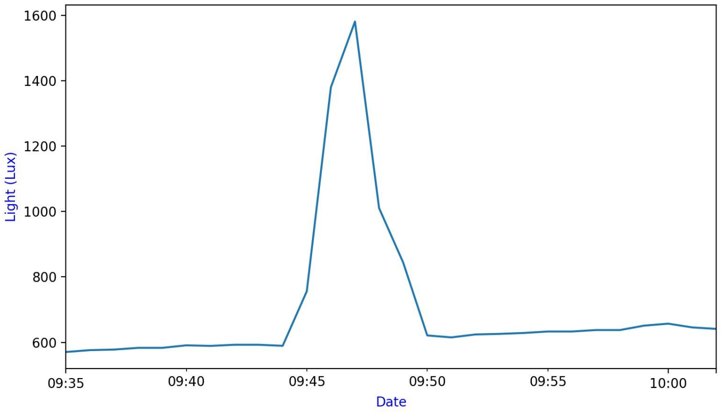
Light measurements on Saturday 2015-02-07 between 9:35am and 9:52am

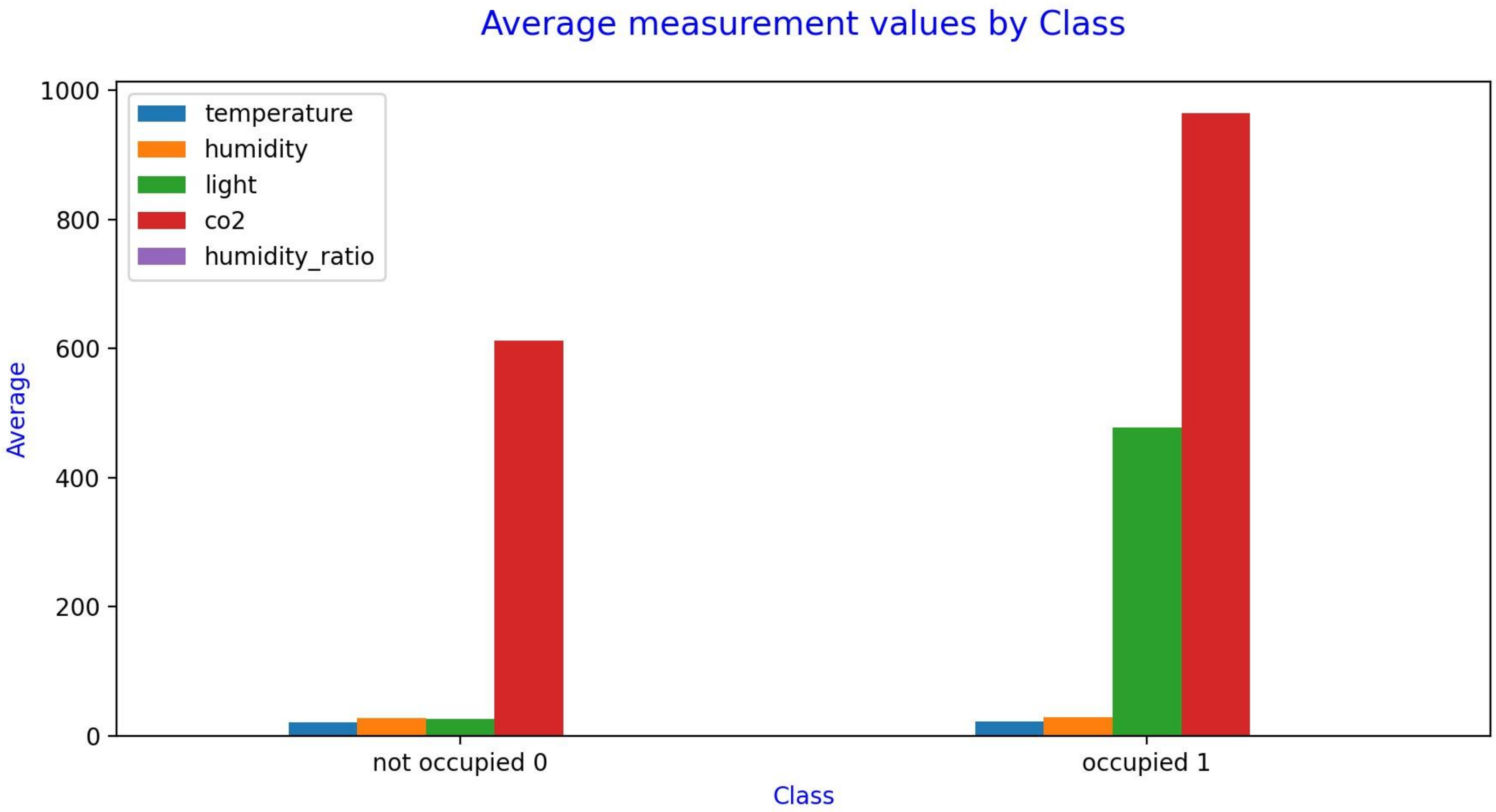


Light measurements on Thursday 2015-02-12

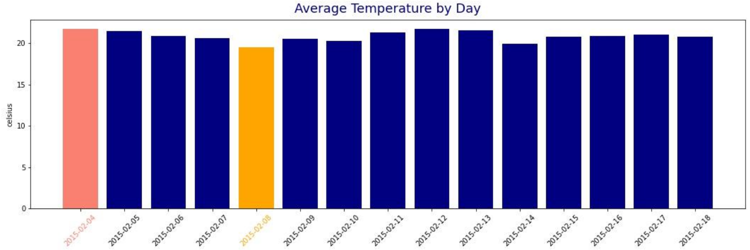


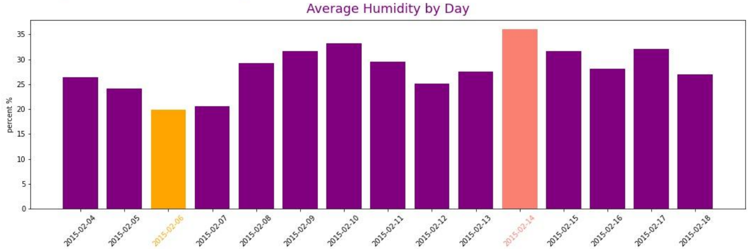
Light measurements on Thursday 2015-02-12 9:35am to 10:02am

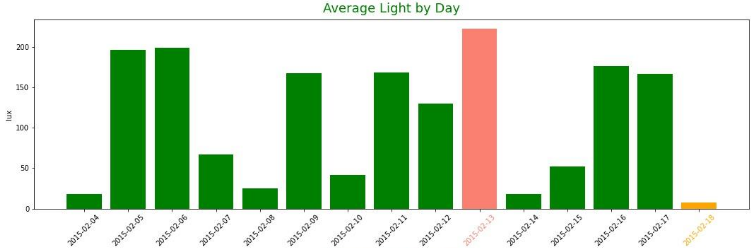


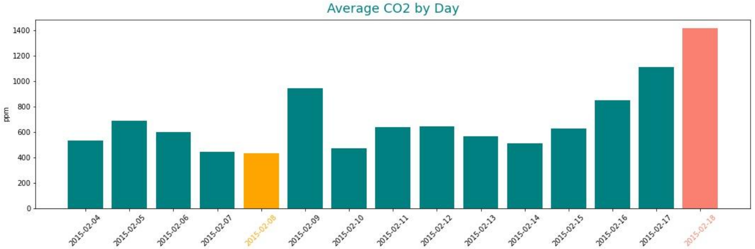


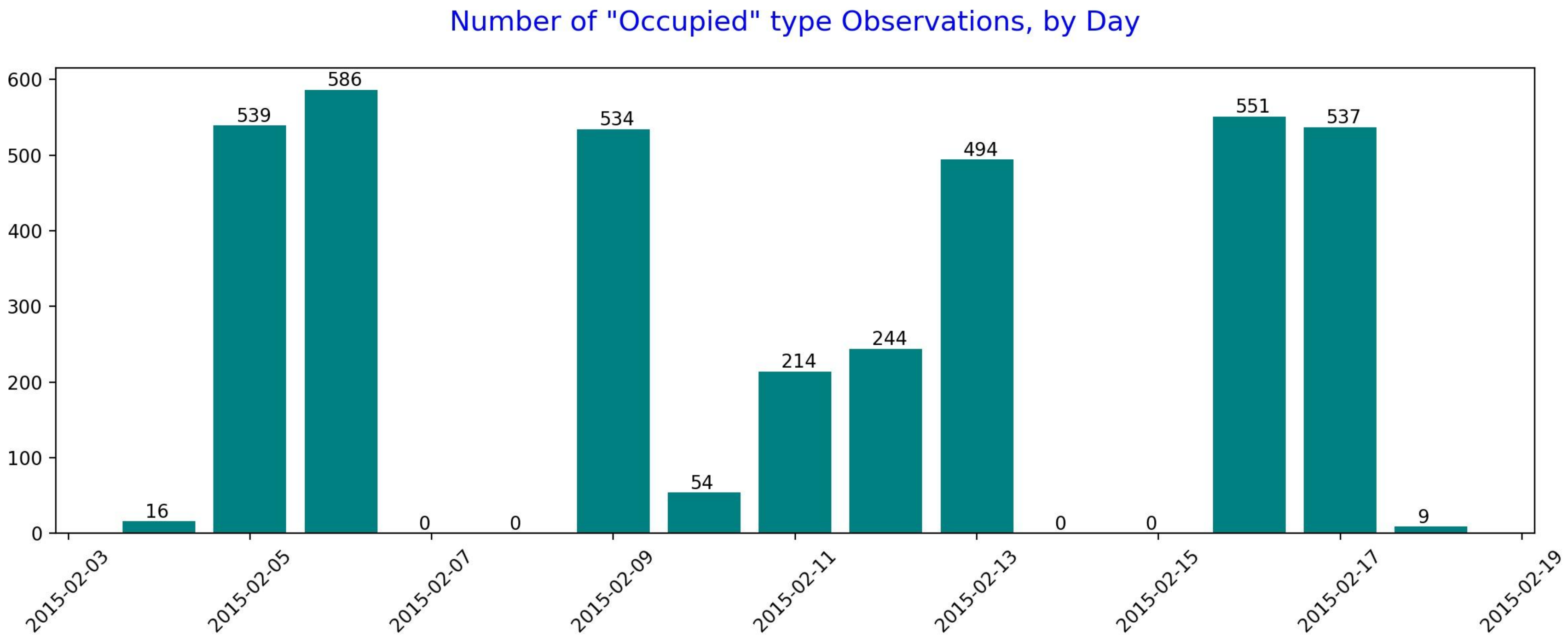
Average measurement values by Class
temperature
humidity
light
co2
humidity_ratio
1000
800
600
400
200
0
Average
not occupied 0
occupied 1
Class

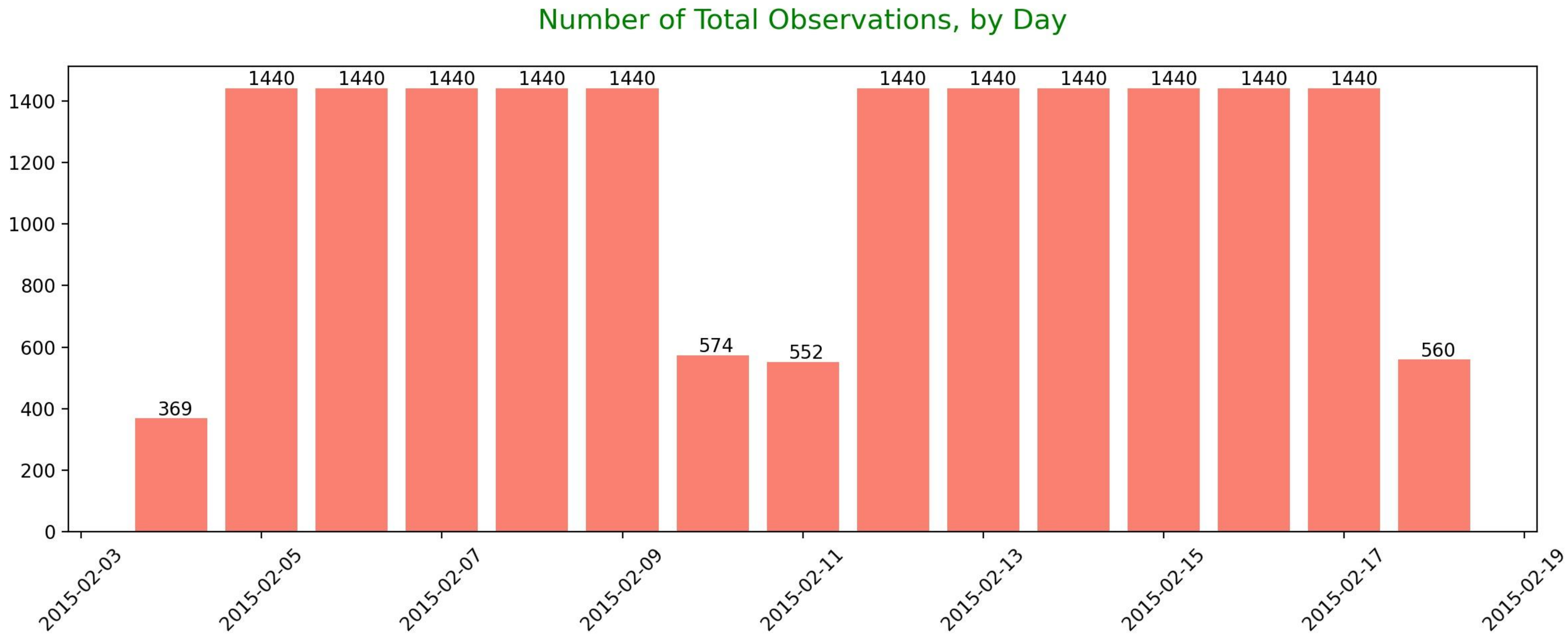

Average Temperature by Day
celsius
0
5
10
15
20
2015-02-04
2015-02-05
2015-02-06
2015-02-07
2015-02-08
2015-02-09
2015-02-10
2015-02-11
2015-02-12
2015-02-13
2015-02-14
2015-02-15
2015-02-16
2015-02-17
2015-02-18


Average Humidity by Day
percent %
0
5
10
15
20
25
30
35
2015-02-04
2015-02-05
2015-02-06
2015-02-07
2015-02-08
2015-02-09
2015-02-10
2015-02-11
2015-02-12
2015-02-13
2015-02-14
2015-02-15
2015-02-16
2015-02-17
2015-02-18


Average Light by Day
lux
0
50
100
150
200
2015-02-04
2015-02-05
2015-02-06
2015-02-07
2015-02-08
2015-02-09
2015-02-10
2015-02-11
2015-02-12
2015-02-13
2015-02-14
2015-02-15
2015-02-16
2015-02-17
2015-02-18

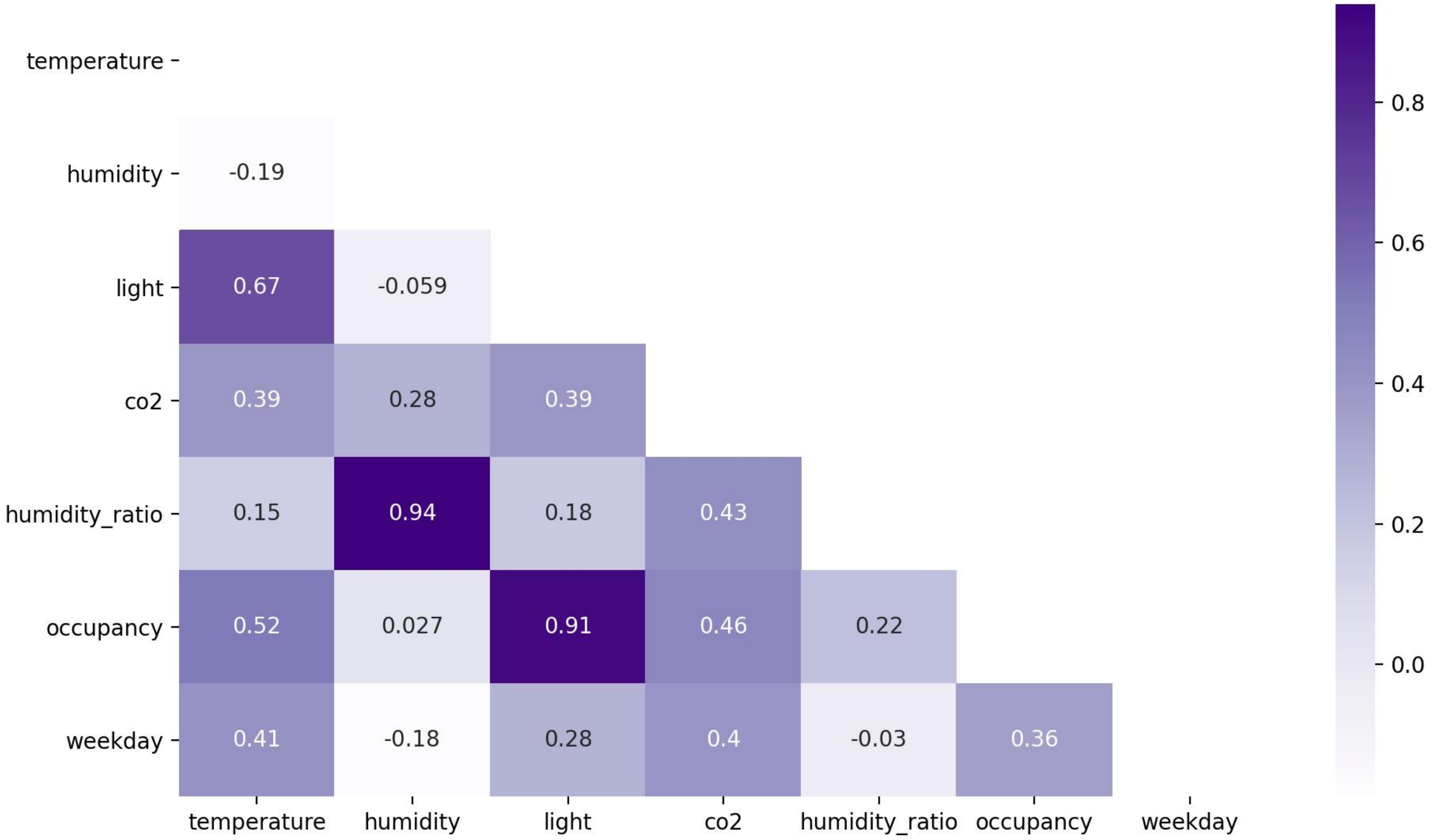

Average CO2 by Day
ppm
0
200
400
600
800
1000
1200
1400
2015-02-04
2015-02-05
2015-02-06
2015-02-07
2015-02-08
2015-02-09
2015-02-10
2015-02-11
2015-02-12
2015-02-13
2015-02-14
2015-02-15
2015-02-16
2015-02-17
2015-02-18


Number of "Occupied" type Observations, by Day
600
500
400
300
200
100
0
16
539
586
0
0
534
54
214
244
494
0
0
551
537
9
2015-02-03
2015-02-05
2015-02-07
2015-02-09
2015-02-11
2015-02-13
2015-02-15
2015-02-17
2015-02-19



Number of Total Observations, by Day
1440
1440
1440
1440
1440
1440
1440
1440
1440
1440
1440
369
574
552
560
1400
1200
1000
800
600
400
200
0
2015-02-03
2015-02-05
2015-02-07
2015-02-09
2015-02-11
2015-02-13
2015-02-15
2015-02-17
2015-02-19

Correlation Observed between Occupancy and Measurements



Modeling

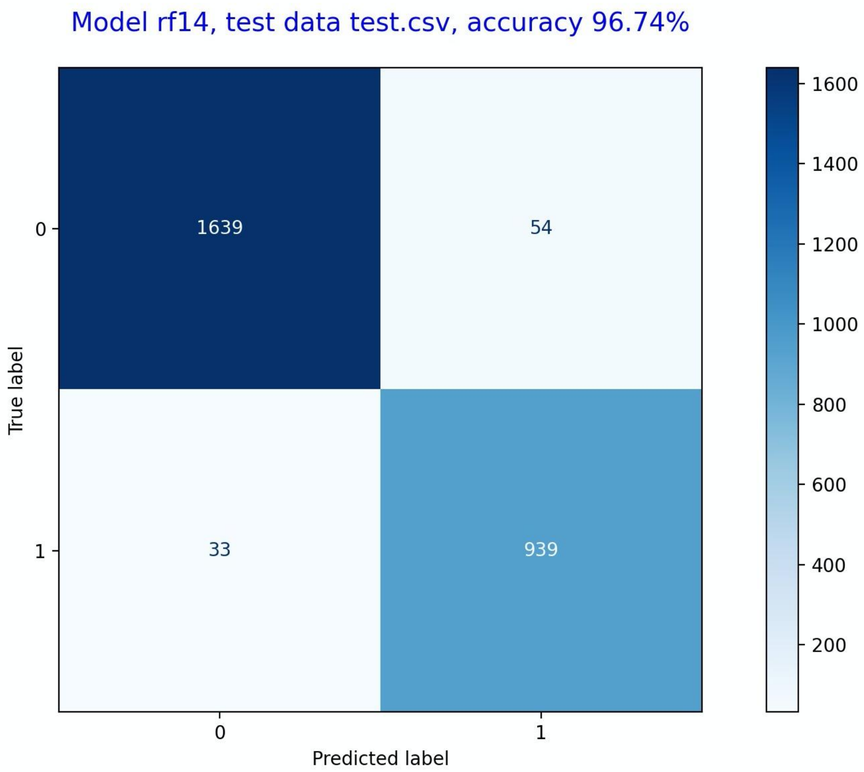
- Models were trained on 75% of the combined data, with 25% left for testing.
- 63 models were run on the [temperature, light, co2, humidity, humidity_ratio] features..
- 154 models were run on the [temperature, light, co2, humidity, humidity_ratio, weekday] features.
- A total of 217 models were run.
- The best model was Random Forest. At training time, the model reported 99.60% accuracy on a 75/25 train/test split of combined data, 96.74% accuracy on test data with the door open, and 99.72% accuracy on test data with the door closed.

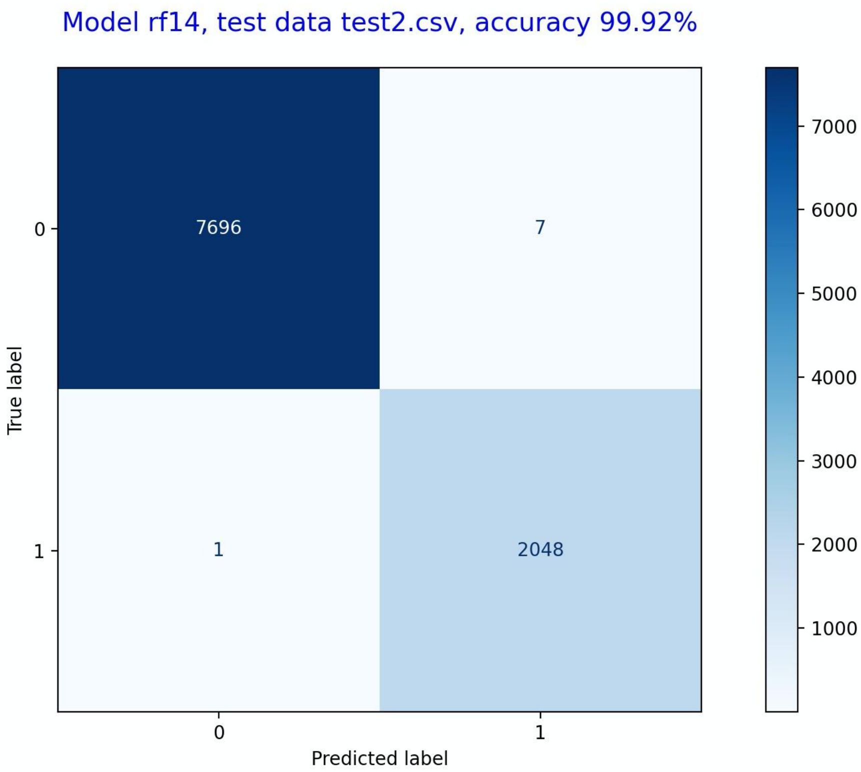
Abbreviation	Meaning
RF	Random Forest
LDA	Linear Discriminant Analysis
GBM	Gradient Boosting Machine
AdaBoost	Adaptive Boosting
KNN	K-Nearest Neighbor
SVC	Support Vector Classification

Model Id	Model name	Features	Best score	Train score	Test score	Sensitivity	Specificity	Precision	Accuracy	F1-score
14	**rf14**	temperature, humidity, light, humidity_ratio, weekday	0.994	0.9996	0.996	0.9905	0.9974	0.9905	0.996	0.9905
12	**rf12**	temperature, humidity, light, co2, weekday	0.9945	1	0.9958	0.9894	0.9974	0.9905	0.9958	0.9899
11	**rf11**	temperature, humidity, light, co2, humidity_ratio, weekday	0.9946	1	0.9955	0.9905	0.9969	0.9884	0.9955	0.9894
1	**rf1**	temperature, humidity, light, co2	0.994	1	0.9955	0.9884	0.9974	0.9905	0.9955	0.9894
0	**rf0**	temperature, humidity, light, co2, humidity_ratio	0.9938	1	0.9953	0.9873	0.9974	0.9904	0.9953	0.9889
77	**gbm11**	temperature, humidity, light, co2, humidity_ratio, weekday	0.994	1	0.9951	0.9884	0.9969	0.9884	0.9951	0.9884
67	**gbm1**	temperature, humidity, light, co2	0.9932	0.9989	0.9949	0.9873	0.9969	0.9883	0.9949	0.9878
18	**rf18**	humidity, light, weekday	0.992	0.9994	0.9949	0.9884	0.9966	0.9873	0.9949	0.9878
3	**rf3**	temperature, humidity, light, humidity_ratio	0.9931	0.9996	0.9944	0.9831	0.9974	0.9904	0.9944	0.9867
66	**gbm0**	temperature, humidity, light, co2, humidity_ratio	0.993	1	0.9944	0.9862	0.9966	0.9873	0.9944	0.9868

Random Forest

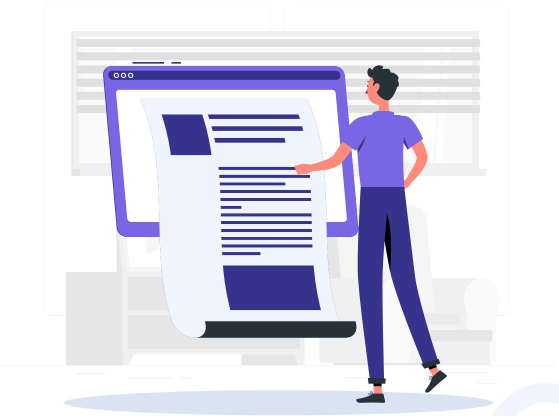
temperature, humidity, light, humidity_ratio, weekday





Conclusions

Recommendations



Next Steps

Resources

(1) Sensitivity and Specificity
(2) Accuracy
(3) Precision
(4) Executive Summary
(5) How to Predict Room Occupancy Based on Environmental Factors
(6) Technical Report and Project
(7) Improving Prediction of Office Room Occupancy Through Random Sampling
(8) Occupancy
(9) Room Occupancy Detection
(10) Classroom Occupancy Project
(11) Accurate occupancy detection of an office room from light, temperature, humidity and CO2 measurements using statistical learning models. Luis M. Candanedo, VÃ©ronique Feldheim. Energy and Buildings. Volume 112, 15 January 2016, Pages 28-39.
(12) Calculate day in the past
(13) V.L.Erickson, M.Á.Carreira-Perpinán, A.E.Cerpa, OBSERVE:Occupancy-based system for efficient reduction of HVAC energy, in: Proceedings of the 10th International Conference on, IEEE, Information Processing in Sensor Networks (IPSN), Chicago, IL, 2011, pp. 258–269.
(14) V.L.Erickson, M.Á.Carreira-Perpinán, A.E.Cerpa,Occupancy modeling and prediction for building energy management, ACM Trans. Sensor Netw. (TOSN) 10 (3) (2014) 42.
(14) Dong B., Andrews B., (2009). Sensor-based occupancy behavioral pattern recognition for energy and comfort management in intelligent buildings. Proceedings of Building Simulation.
(15) J. Brooks, S. Goyal, R. Subramany, Y. Lin, T. Middelkoop, L. Arpan, L. Carloni, P. Barooah, An experimental investigation of occupancy-based energy-efficient control of commercial building indoor climate, in: Proceeding of the IEEE 53rd Annual Conference on, IEEE, Decision and Control (CDC), Los Angeles, CA, 2014, pp. 5680–5685.
(16) J. Brooks, S. Kumar, S. Goyal, R. Subramany, P. Barooah, Energy-efficient control of under-actuated HVAC zones in commercial buildings, Energy Build. 93 (2015) 160–168.

Title

ALTERNATIVE RESOURCES

Find more illustrations like these on **Stories by Freepik**