

# TIMEPAC Academy

## Session 3

### Data extraction from the multiple sources

Presenter: Álvaro Sicilia (La Salle-URL)

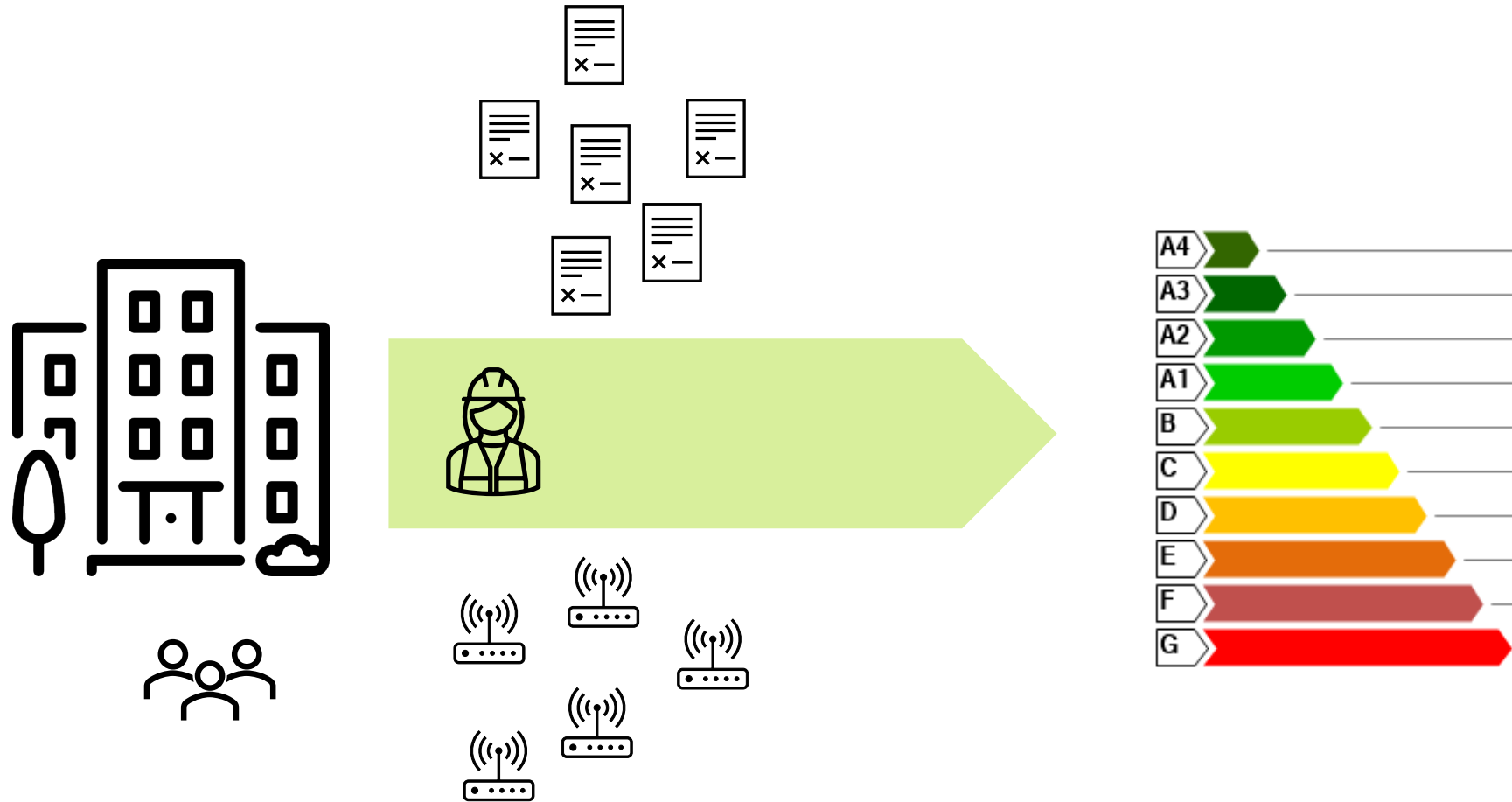
29 February 2024



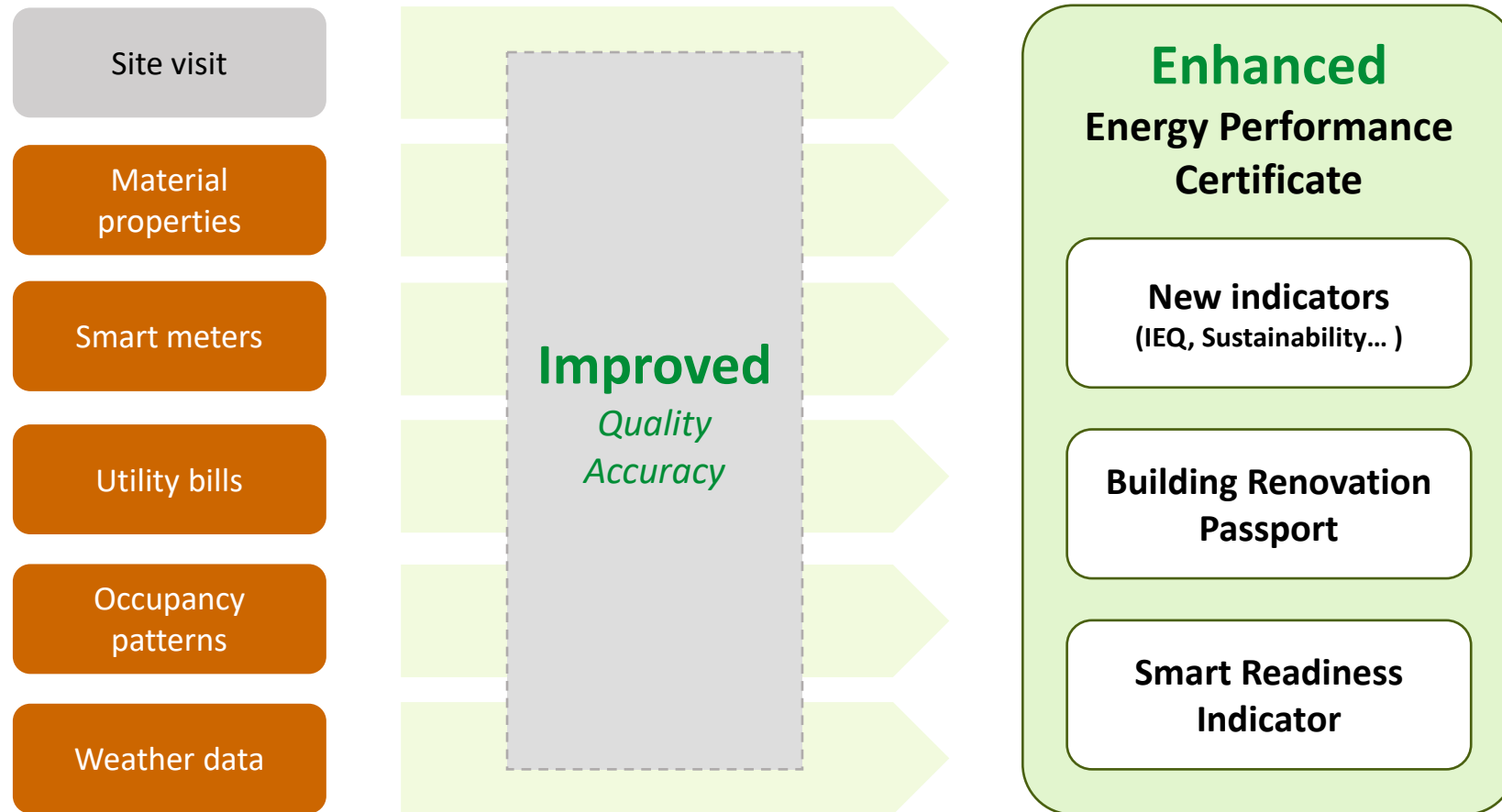
Politecnico di Torino



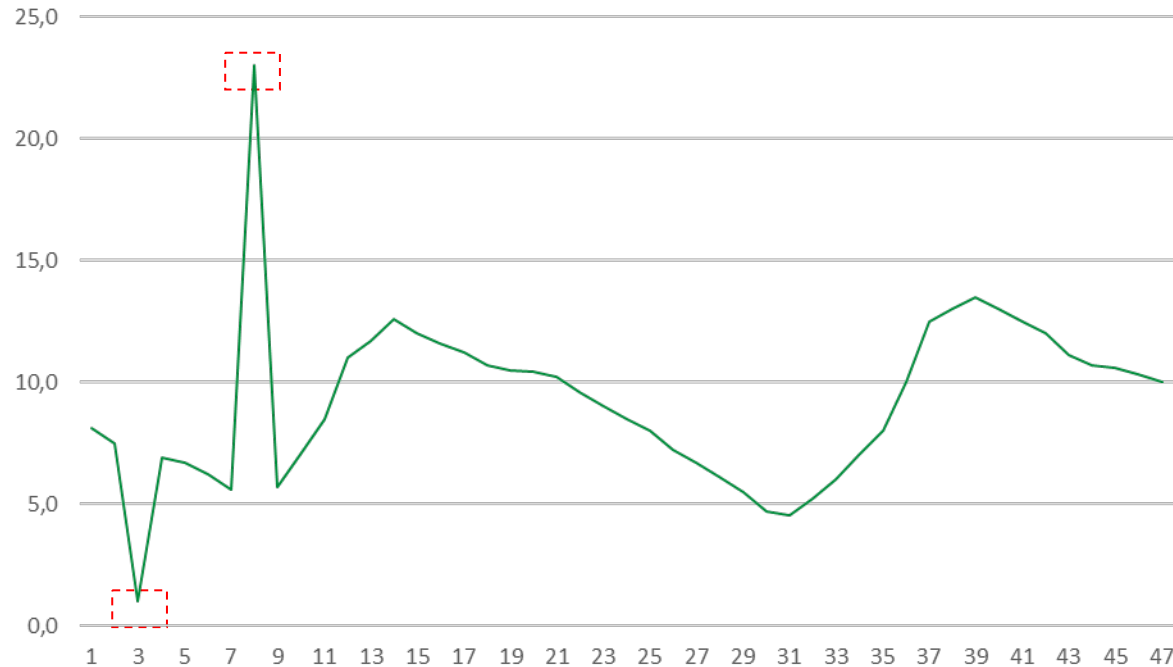
# Data extraction from the multiple sources



# Data extraction from the multiple sources



# Handling data issues: Outliers

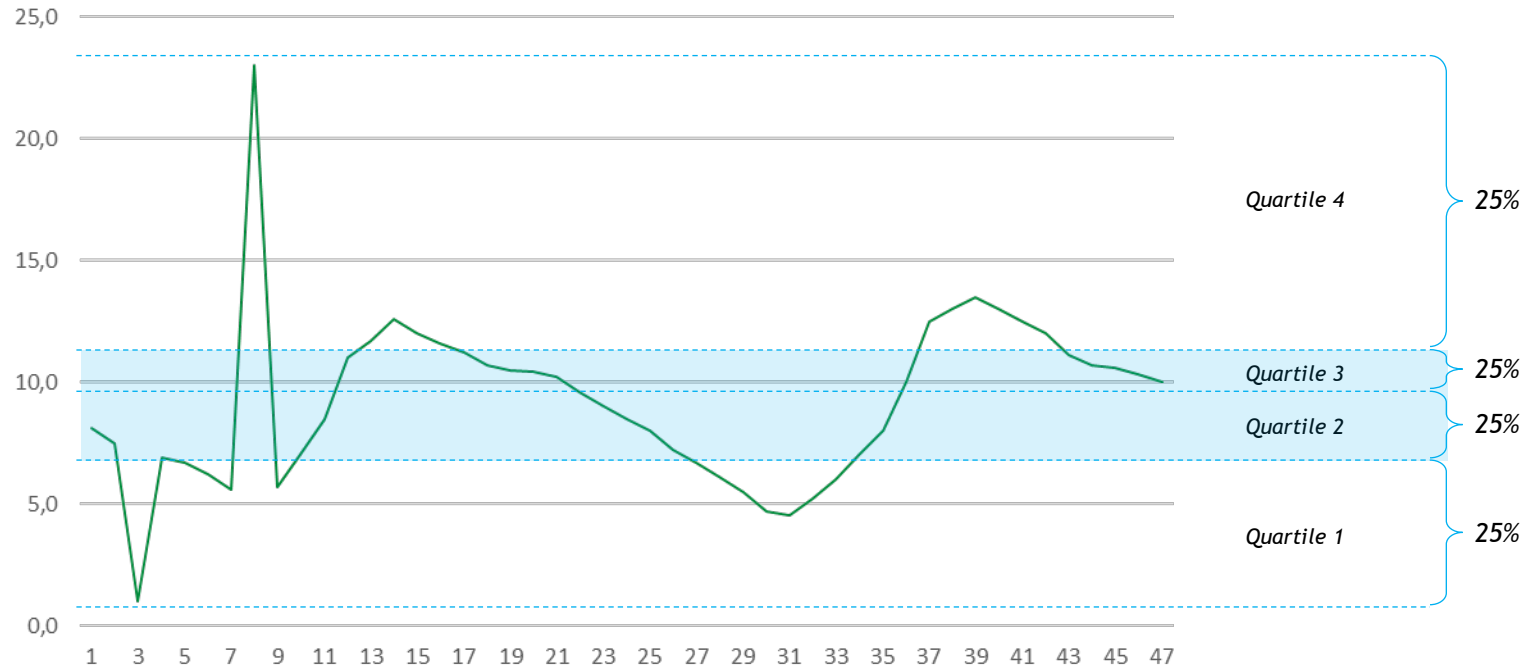


## Data inconsistency:

- An outlier is a data point that falls at an abnormal distance from the other values.
- Outliers can indicate potential errors, anomalies, or irregularities.
- Problems with the equipment (i.e., sensors)
- Energy simulation tools may be sensitive to outliers leading to bad simulations.

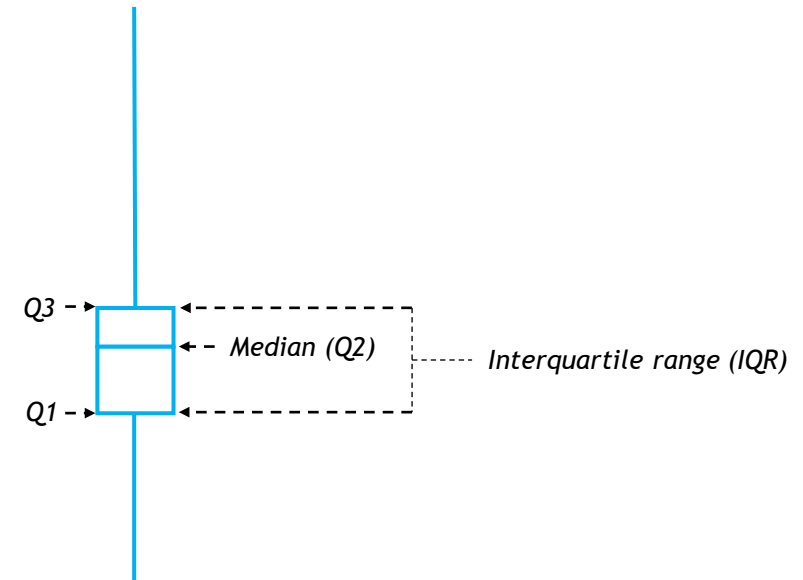
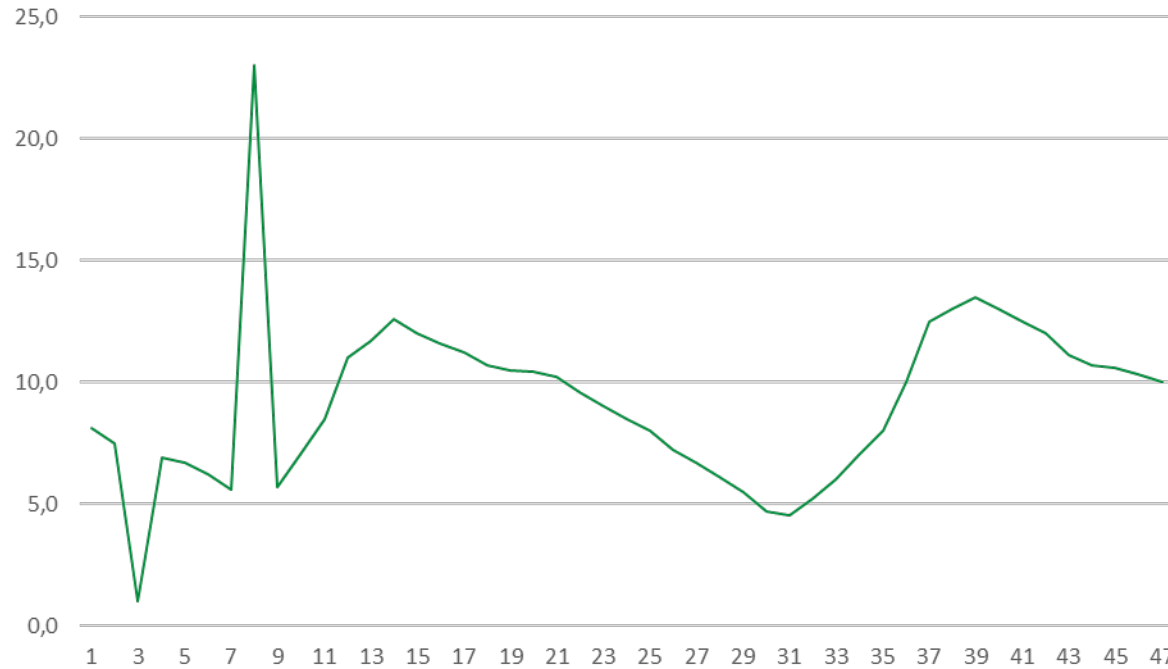
# Handling data issues: Outliers

## Inter Quartile Range (IQR) method



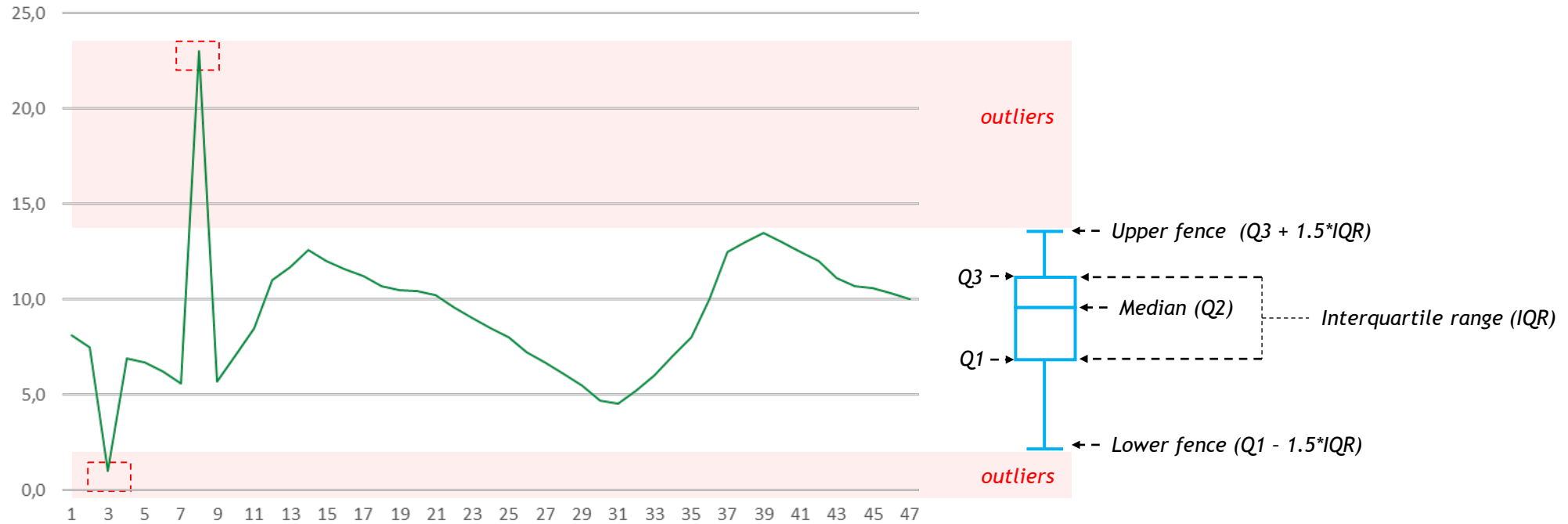
# Handling data issues: Outliers

## Inter Quartile Range (IQR) method



# Handling data issues: Outliers

## Inter Quartile Range (IQR) method



# Practical example

## Indoor Environmental Quality calculation (with Excel)

	A	B	C	D	E	F	G	H	I	J	K	L
1	Date/Time			Boundary conditions		Comfort temperatures				Occupancy hours	Space to be evaluated	
2	Month	Day	Hour	Ext. air temp.	Running mean out. temp.	Comfort temp.	Upper limit temp.	Lower limit temp.				Ind. operative temp.
3				$\theta_{ext}$	$\theta_{rm}$	$\theta_c$	$\theta_{c,upper}$	$\theta_{c,lower}$			$\theta_{int,op}$	
4				[°C]	[°C]	[°C]	[°C]	[°C]		[-]	[°C]	
5	1	1	0	8,1	10,7	22,3	26,3	17,3		0	21,6	
6	1	1	1	7,5	10,7	22,3	26,3	17,3		0	21,6	
7	1	1	2	1,0	10,7	22,3	26,3	17,3		0	21,5	
8	1	1	3	6,9	10,7	22,3	26,3	17,3		0	21,4	
9	1	1	4	6,7	10,7	22,3	26,3	17,3		0	10,0	
10	1	1	5	6,2	10,7	22,3	26,3	17,3		0	21,1	
11	1	1	6	5,6	10,7	22,3	26,3	17,3		0	21,0	
12	1	1	7	23,0	10,7	22,3	26,3	17,3		0	21,3	
13	1	1	8	5,7	10,7	22,3	26,3	17,3		1	21,4	
14	1	1	9	7,1	10,7	22,3	26,3	17,3		1	21,8	
15	1	1	10	8,5	10,7	22,3	26,3	17,3		1	22,6	
16	1	1	11	11,0	10,7	22,3	26,3	17,3		1	23,3	
17	1	1	12	11,7	10,7	22,3	26,3	17,3		1	23,8	
18	1	1	13	12,6	10,7	22,3	26,3	17,3		1	24,2	
19	1	1	14	12,0	10,7	22,3	26,3	17,3		1	35,0	
20	1	1	15	11,6	10,7	22,3	26,3	17,3		1	24,0	
21	1	1	16	11,2	10,7	22,3	26,3	17,3		1	23,5	
22	1	1	17	10,7	10,7	22,3	26,3	17,3		1	23,1	
23	1	1	18	10,5	10,7	22,3	26,3	17,3		0	22,9	
24	1	1	19	10,4	10,7	22,3	26,3	17,3		0	22,7	
25	1	1	20	10,2	10,7	22,3	26,3	17,3		0	22,5	
26	1	1	21	9,6	10,7	22,3	26,3	17,3		0	22,2	
27	1	1	22	9,0	10,7	22,3	26,3	17,3		0	22,1	
28	1	1	23	8,5	10,7	22,3	26,3	17,3		0	21,9	
29	1	2	0	8,0	10,1	22,1	26,1	17,1		0	21,7	
30	1	2	1	7,2	10,1	22,1	26,1	17,1		0	21,6	
31	1	2	2	6,7	10,1	22,1	26,1	17,1		0	21,5	



# Practical example

## Indoor Environmental Quality calculation (with Excel)

=QUARTILE(D5:D168, 2) //median  
 =QUARTILE(D5:D168, 3) //Q3  
 =QUARTILE(D5:D168, 1) //Q1  
 =QUARTILE(D5:D168, 3) - QUARTILE(D5:D168, 1) // IQR

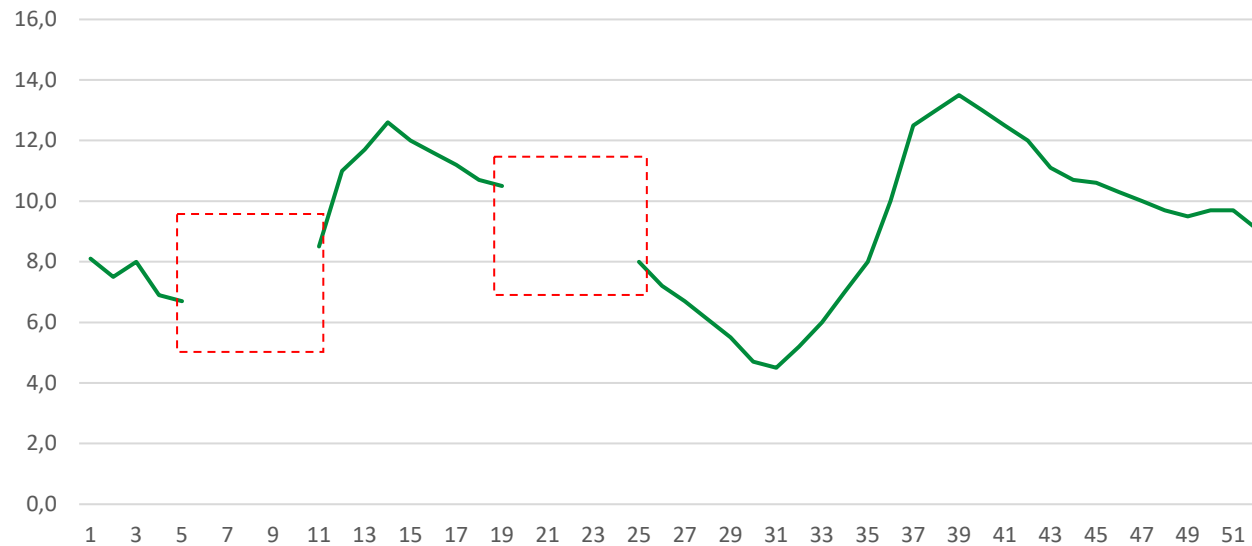
	Date/Time			Boundary conditions		Comfort temperatures			Occupancy hours	Space to be evaluated Ind. operative temp. $\theta_{int.op}$
	Month	Day	Hour	Ext. air temp.	Running mean out. temp.	Comfort temp.	Upper limit temp.	Lower limit temp.		
				$\theta_{ext}$	$\theta_{rm}$	$\theta_c$	$\theta_{c,upper}$	$\theta_{c,lower}$		
				[°C]	[°C]	[°C]	[°C]	[°C]		
1	1	1	0	8,1	10,7	22,3	26,3	17,3	0	21,6
2	1	1	1	7,5	10,7	22,3	26,3	17,3	0	21,6
3	1	1	2	1,0	10,7	22,3	26,3	17,3	0	21,5
4	1	1	3	6,9	10,7	22,3	26,3	17,3	0	21,4
5	1	1	4	6,7	10,7	22,3	26,3	17,3	0	21,1
6	1	1	5	6,2	10,7	22,3	26,3	17,3	0	21,0
7	1	1	6	5,6	10,7	22,3	26,3	17,3	0	21,3
8	1	1	7	23,0	10,7	22,3	26,3	17,3	0	21,4
9	1	1	8	5,7	10,7	22,3	26,3	17,3	1	21,8
10	1	1	9	7,1	10,7	22,3	26,3	17,3	1	22,6
11	1	1	10	8,5	10,7	22,3	26,3	17,3	1	23,3
12	1	1	11	11,0	10,7	22,3	26,3	17,3	1	23,8
13	1	1	12	11,7	10,7	22,3	26,3	17,3	1	24,2
14	1	1	13	12,6	10,7	22,3	26,3	17,3	1	35,0
15	1	1	14	12,0	10,7	22,3	26,3	17,3	1	24,0
16	1	1	15	11,6	10,7	22,3	26,3	17,3	1	23,5
17	1	1	16	11,2	10,7	22,3	26,3	17,3	1	23,1
18	1	1	17	10,7	10,7	22,3	26,3	17,3	0	22,9
19	1	1	18	10,5	10,7	22,3	26,3	17,3	0	22,7
20	1	1	19	10,4	10,7	22,3	26,3	17,3	0	22,5
21	1	1	20	10,2	10,7	22,3	26,3	17,3	0	22,2
22	1	1	21	9,6	10,7	22,3	26,3	17,3	0	22,1
23	1	1	22	9,0	10,7	22,3	26,3	17,3	0	21,9
24	1	1	23	8,5	10,7	22,3	26,3	17,3	0	21,7
25	1	2	0	8,0	10,1	22,1	26,1	17,1	0	21,6
26	1	2	1	7,2	10,1	22,1	26,1	17,1	0	21,5
27	1	2	2	6,7	10,1	22,1	26,1	17,1	0	21,5

Ext. air temp.	Ind. operative temp.
----------------	----------------------

	Median:	9,7		Median:	22,0
	Third quartile	11,8		Third quartile	23,1
outlier: too small	First quartile	7,6		First quartile	21,4
	Inter quartile range	4,2		Inter quartile range	1,7
			outlier: too small		
outlier too big					
			outlier too big		

=IF(D5 < (P7 - 1.5\*P8),"outlier: too small",  
 IF(D5 > (P6 + 1.5\*P8),"outlier too big", ""))

# Handling data issues: missing points

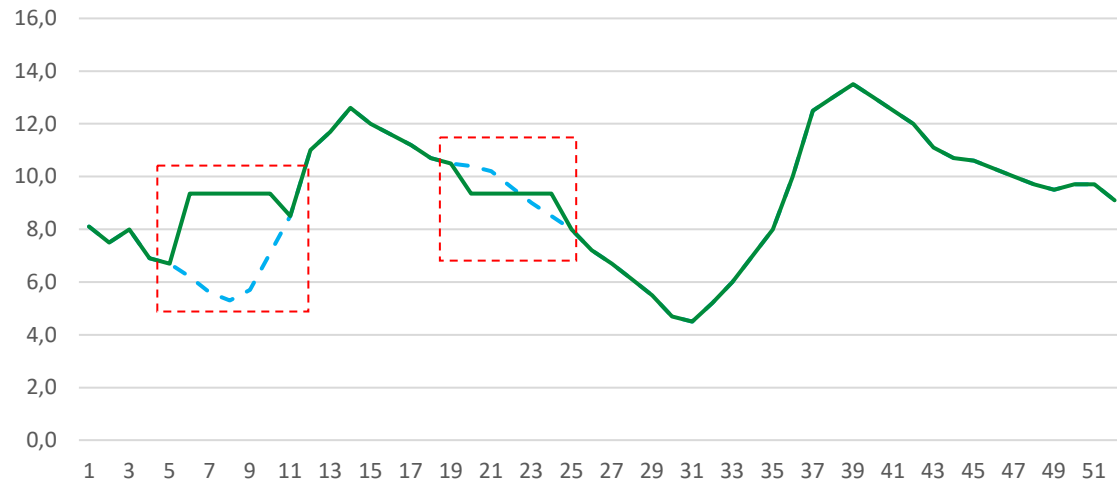


## Incomplete datasets:

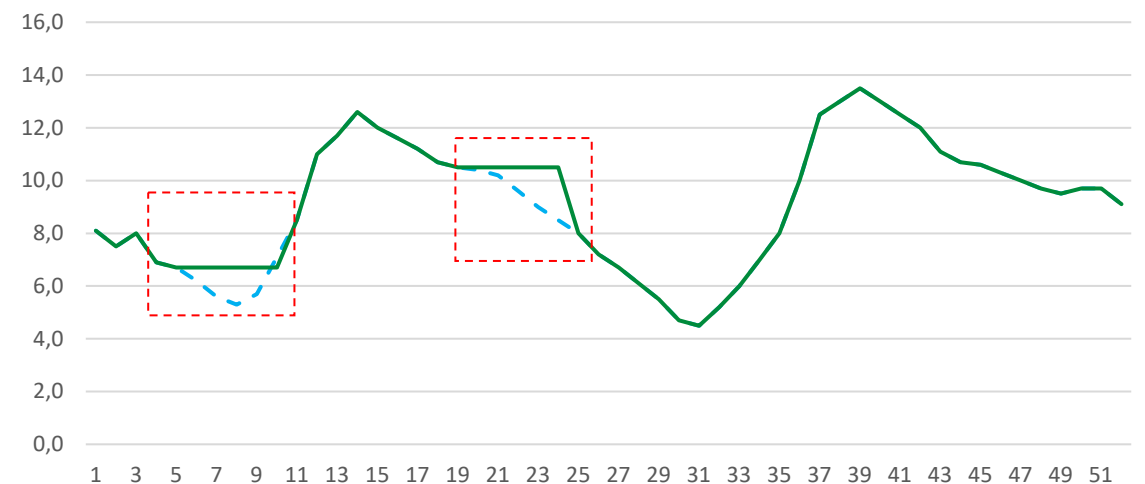
- Incomplete datasets may arise due to gaps in utility records or limitations in monitoring systems
- Missing information can impact the accuracy of EPCs

# Handling data issues: missing points

## 1. Fill with Average/Mean



## 2. Fill with Previous/Next Value:



# Handling data issues: missing points

## Indoor Environmental Quality calculation (with Excel)

1	Date/Time			Boundary conditions		Comfort temperatures			Occupancy hours	Space to be evaluated Ind. operative temp. $\theta_{int,op}$
	Month	Day	Hour	Ext. air temp.	Running mean out. temp.	Comfort temp.	Upper limit temp.	Lower limit temp.		
				$\theta_{ext}$	$\theta_{rm}$	$\theta_c$	$\theta_{c,upper}$	$\theta_{c,lower}$		
				[°C]	[°C]	[°C]	[°C]	[°C]		
2	1	1	0	8,1	10,7	22,3	26,3	17,3	0	21,6
3	1	1	1	7,5	10,7	22,3	26,3	17,3	0	21,6
4	1	1	2	1,0	10,7	22,3	26,3	17,3	0	21,5
5	1	1	3	6,9	10,7	22,3	26,3	17,3	0	21,4
6	1	1	4	6,7	10,7	22,3	26,3	17,3	0	10,0
7	1	1	5		10,7	22,3	26,3	17,3	0	21,1
8	1	1	6		10,7	22,3	26,3	17,3	0	21,0
9	1	1	7		10,7	22,3	26,3	17,3	0	21,3
10	1	1	8		10,7	22,3	26,3	17,3	1	21,4
11	1	1	9		10,7	22,3	26,3	17,3	1	21,8
12	1	1	10	8,5	10,7	22,3	26,3	17,3	1	22,6
13	1	1	11	11,0	10,7	22,3	26,3	17,3	1	23,3
14	1	1	12	11,7	10,7	22,3	26,3	17,3	1	23,8
15	1	1	13	12,6	10,7	22,3	26,3	17,3	1	24,2
16	1	1	14	12,0	10,7	22,3	26,3	17,3	1	35,0
17	1	1	15	11,6	10,7	22,3	26,3	17,3	1	24,0
18	1	1	16	11,2	10,7	22,3	26,3	17,3	1	23,5
19	1	1	17	10,7	10,7	22,3	26,3	17,3	1	23,1
20	1	1	18	10,5	10,7	22,3	26,3	17,3	0	22,9
21	1	1	19		10,7	22,3	26,3	17,3	0	22,7
22	1	1	20		10,7	22,3	26,3	17,3	0	22,5
23	1	1	21		10,7	22,3	26,3	17,3	0	22,2
24	1	1	22		10,7	22,3	26,3	17,3	0	22,1
25	1	1	23		10,7	22,3	26,3	17,3	0	21,9
26	1	2	0	8,0	10,1	22,1	26,1	17,1	0	21,7
27	1	2	1	7,2	10,1	22,1	26,1	17,1	0	21,6
28	1	2	2	6,7	10,1	22,1	26,1	17,1	0	21,5

Average: 9,2

=AVERAGE(D5:D168)

=IF(ISBLANK(D10), D9, D10)

# Take aways

The new EPBD will require to work with more datasets.

Do not take for granted the accuracy and reliability of data  
wherever it comes from!

Validate and fix the data you are using for your EPC generation.

**If you would like more information,  
please visit [www.timepac.eu](http://www.timepac.eu) or contact us at  
[alvaro.sicilia@salle.url.edu](mailto:alvaro.sicilia@salle.url.edu)**

Thanks for your attention!