

## extract\_data1

May 18, 2025

```
[2]: import pandas as pd
import numpy as np
```

```
[3]: path_file1 = '/home/darkcover/Documentos/Gan/Data/Real/trainingData.csv'
data1 = pd.read_csv(path_file1)

data1.head()
```

```
[3]:
```

	WAP001	WAP002	WAP003	WAP004	WAP005	WAP006	WAP007	WAP008	WAP009	\
0	100	100	100	100	100	100	100	100	100	
1	100	100	100	100	100	100	100	100	100	
2	100	100	100	100	100	100	100	-97	100	
3	100	100	100	100	100	100	100	100	100	
4	100	100	100	100	100	100	100	100	100	

  

	WAP010	...	WAP520	LONGITUDE	LATITUDE	FLOOR	BUILDINGID	SPACEID	\
0	100	...	100	-7541.2643	4.864921e+06	2	1	106	
1	100	...	100	-7536.6212	4.864934e+06	2	1	106	
2	100	...	100	-7519.1524	4.864950e+06	2	1	103	
3	100	...	100	-7524.5704	4.864934e+06	2	1	102	
4	100	...	100	-7632.1436	4.864982e+06	0	0	122	

  

	RELATIVEPOSITION	USERID	PHONEID	TIMESTAMP
0		2	2	23 1371713733
1		2	2	23 1371713691
2		2	2	23 1371714095
3		2	2	23 1371713807
4		2	11	13 1369909710

[5 rows x 529 columns]

```
[4]: data1.describe()
```

```
[4]:
```

	WAP001	WAP002	WAP003	WAP004	WAP005	\
count	19937.000000	19937.000000	19937.0	19937.0	19937.000000	
mean	99.823644	99.820936	100.0	100.0	99.613733	
std	5.866842	5.798156	0.0	0.0	8.615657	
min	-97.000000	-90.000000	100.0	100.0	-97.000000	

25%	100.000000	100.000000	100.0	100.0	100.000000
50%	100.000000	100.000000	100.0	100.0	100.000000
75%	100.000000	100.000000	100.0	100.0	100.000000
max	100.000000	100.000000	100.0	100.0	100.000000

	WAP006	WAP007	WAP008	WAP009	WAP010 \
count	19937.000000	19937.000000	19937.000000	19937.000000	19937.000000
mean	97.130461	94.733661	93.820234	94.693936	99.163766
std	22.931890	30.541335	33.010404	30.305084	12.634045
min	-98.000000	-99.000000	-98.000000	-98.000000	-99.000000
25%	100.000000	100.000000	100.000000	100.000000	100.000000
50%	100.000000	100.000000	100.000000	100.000000	100.000000
75%	100.000000	100.000000	100.000000	100.000000	100.000000
max	100.000000	100.000000	100.000000	100.000000	100.000000

	...	WAP520	LONGITUDE	LATITUDE	FLOOR	BUILDINGID \
count	...	19937.0	19937.000000	1.993700e+04	19937.000000	19937.000000
mean	...	100.0	-7464.275947	4.864871e+06	1.674575	1.212820
std	...	0.0	123.402010	6.693318e+01	1.223078	0.833139
min	...	100.0	-7691.338400	4.864746e+06	0.000000	0.000000
25%	...	100.0	-7594.737000	4.864821e+06	1.000000	0.000000
50%	...	100.0	-7423.060900	4.864852e+06	2.000000	1.000000
75%	...	100.0	-7359.193000	4.864930e+06	3.000000	2.000000
max	...	100.0	-7300.818990	4.865017e+06	4.000000	2.000000

	SPACEID	RELATIVEPOSITION	USERID	PHONEID \
count	19937.000000	19937.000000	19937.000000	19937.000000
mean	148.429954	1.833024	9.068014	13.021869
std	58.342106	0.372964	4.988720	5.362410
min	1.000000	1.000000	1.000000	1.000000
25%	110.000000	2.000000	5.000000	8.000000
50%	129.000000	2.000000	11.000000	13.000000
75%	207.000000	2.000000	13.000000	14.000000
max	254.000000	2.000000	18.000000	24.000000

	TIMESTAMP
count	1.993700e+04
mean	1.371421e+09
std	5.572054e+05
min	1.369909e+09
25%	1.371056e+09
50%	1.371716e+09
75%	1.371721e+09
max	1.371738e+09

[8 rows x 529 columns]

```
[5]: data1.columns
```

```
[5]: Index(['WAP001', 'WAP002', 'WAP003', 'WAP004', 'WAP005', 'WAP006', 'WAP007',  
         'WAP008', 'WAP009', 'WAP010',  
         ...  
         'WAP520', 'LONGITUDE', 'LATITUDE', 'FLOOR', 'BUILDINGID', 'SPACEID',  
         'RELATIVEPOSITION', 'USERID', 'PHONEID', 'TIMESTAMP'],  
        dtype='object', length=529)
```

```
[6]: data2 = data1[['LONGITUDE', 'LATITUDE', 'FLOOR', 'BUILDINGID']]  
data2.head()
```

```
[6]:
```

	LONGITUDE	LATITUDE	FLOOR	BUILDINGID
0	-7541.2643	4.864921e+06	2	1
1	-7536.6212	4.864934e+06	2	1
2	-7519.1524	4.864950e+06	2	1
3	-7524.5704	4.864934e+06	2	1
4	-7632.1436	4.864982e+06	0	0

```
[7]: data2.describe()
```

```
[7]:
```

	LONGITUDE	LATITUDE	FLOOR	BUILDINGID
count	19937.000000	1.993700e+04	19937.000000	19937.000000
mean	-7464.275947	4.864871e+06	1.674575	1.212820
std	123.402010	6.693318e+01	1.223078	0.833139
min	-7691.338400	4.864746e+06	0.000000	0.000000
25%	-7594.737000	4.864821e+06	1.000000	0.000000
50%	-7423.060900	4.864852e+06	2.000000	1.000000
75%	-7359.193000	4.864930e+06	3.000000	2.000000
max	-7300.818990	4.865017e+06	4.000000	2.000000

A data possui entradas nas colunas com valor igual a 100

11 de mai 09:39

trainingData.csv — LibreOffice Calc

File Edit View Insert Format Styles Sheet Data Tools Window Help

Libertation Sans 10 pt

A1 f. Σ WAP001

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
262	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	-91	-91	100	100	100
263	100	100	100	100	100	-94	100	100	100	100	100	100	100	100	100	100	100	100	100	100
264	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	-95	-96	100	100	100
265	100	100	100	100	100	100	100	100	100	-95	100	100	100	100	100	-93	-96	100	100	100
266	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	-93	-92	100	100	100
267	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
268	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
269	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
270	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
271	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
272	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
273	100	100	100	100	100	100	100	100	100	100	-95	100	100	100	100	100	100	100	100	100
274	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
275	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
276	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
277	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
278	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
279	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
280	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
281	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
282	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
283	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
284	100	100	100	100	100	100	100	100	100	100	-90	100	100	100	100	100	100	100	100	100
285	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
286	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
287	100	100	100	100	100	100	100	100	100	100	100	-87	100	100	100	100	100	100	100	100
288	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
289	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
290	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
291	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
292	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
293	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
294	100	100	100	100	100	100	100	100	100	100	-93	100	100	100	100	100	100	100	100	100
295	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
296	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
297	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
298	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
299	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
300	100	100	100	100	100	100	100	100	100	100	-90	100	100	100	100	100	100	100	100	100
301	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
302	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
303	100	100	100	100	100	100	100	100	100	100	-93	-91	100	100	100	100	100	100	100	100
304	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

trainingData

Find Find All Formatted Display Match Case

Sheet 1 of 1

Default

Portuguese (Brazil)

T

Average: ; Sum: 0

11 de mai 09:39

ujindoorsubset\_building1\_floor2.csv — LibreOffice Calc

File Edit View Insert Format Styles Sheet Data Tools Window Help

LibreOffice Calc

10 pt B Z U A

E5 f. Σ = -110

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
313	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-79	-79	-110	-110	-110
314	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-76	-77	-110	-110	-110
315	-110	-110	-110	-110	-110	-110	-110	-110	-110	-95	-110	-110	-110	-110	-85	-84	-110	-110	-110
316	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-77	-76	-110	-110	-110
317	-110	-110	-110	-110	-110	-110	-110	-110	-110	-99	-110	-110	-110	-110	-87	-88	-110	-110	-110
318	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-90	-90	-110	-110	-110
319	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-84	-84	-110	-110	-110
320	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-87	-87	-110	-110	-110
321	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-92	-93	-110	-110	-110
322	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-79	-79	-110	-110	-110
323	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-85	-80	-110	-110	-110
324	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
325	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-92	-92	-110	-110	-110
326	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-93	-92	-110	-110	-110
327	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-93	-94	-110	-110	-110
328	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-89	-87	-110	-110	-110
329	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-93	-95	-110	-110	-110
330	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-90	-92	-110	-110	-110
331	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
332	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
333	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
334	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
335	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
336	-110	-110	-110	-110	-110	-110	-110	-97	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
337	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
338	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
339	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
340	-110	-110	-110	-110	-110	-110	-110	-95	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
341	-110	-110	-110	-110	-110	-110	-110	-90	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
342	-110	-110	-110	-110	-110	-110	-110	-83	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
343	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-82	-110	-110	-110
344	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
345	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-89	-90	-110	-110	-110
346	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
347	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
348	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-86	-86	-110	-110	-110
349	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-83	-84	-110	-110	-110
350	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
351	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-98	-98	-110	-110	-110
352	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-95	-110	-110	-110	-110
353	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-96	-94	-110	-110	-110
354	-110	-110	-110	-110	-110	-110	-110	-110	-95	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110
355	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-110	-77	-77	-110	-110	-110

ujindoorsubset\_building1\_floor2

Find

Sheet 1 of 1

Default Portuguese (Brazil)

Average: -110; Sum: -110

100%

Utilizaremos os dados enviados como base, ficando o ajuste para mudança de valor. Podemos confirmar ?

```
[8]: # Carregar base
df = pd.read_csv(path_file1)

# Substituir valor ausente 100 por -110
rssi_columns = df.columns[:520]
df[rssi_columns] = df[rssi_columns].replace(100, -110)

# Filtro para prédio 1, andar 2
df_filtered = df[(df['BUILDINGID'] == 1) & (df['FLOOR'] == 2)].copy()

# Selecionar RSSI + coordenadas
X_rssi = df_filtered[rssi_columns].values
y_coords = df_filtered[['LONGITUDE', 'LATITUDE']].values
```

```
[9]: # Salvar RSSI + coordenadas do Building 1 - Floor 2
df_filtered.to_csv("/home/darkcover/Documentos/Gan/Data/
↳ujindoorsubset_building1_floor2.csv", index=False)
```