

Dataset

The dataset contains six files, each providing information on water consumption measurements, hierarchical geography, customer profiles, and meteorological information. You are free to add as many additional data sources as you consider necessary to enrich your model.

- caudales.csv
- gis.csv
- abonado.csv
- cliente.csv
- Contador.csv
- Clima.csv

1. caudales.csv

This file contains information related to water consumption in the city of Villarubia (Service VR) for the time period between August 2021 and December 2022.

These are the variables in caudales.csv dataset:

- **Canonical**: Indicates the type of element from which the measurement is obtained.
- **RowKey**: Date and time when the measurement was made.
- **INF_Label**: Specific information about the canonical.
- **INF_Value**: Value of the measurement.
- **STA_Label**: Identification number of the Neta sector and the canonical.
- **Sector_Neta**: Indicates the sector where the element is located.

Note that only the water supply services of Villarubia need to be considered for analysis and forecasting. Also, pay attention to the frequency of measurements.

Canonical variable		Caudales.csv example					
0	DEPOSITO	Canonical	RowKey	INF_Label	INF_Value	STA_Label	Sector_Neta
1	DEPOSITO_CAUDAL_SALIDA	0	DEPOSITO 2021-08-01T00:15:00.000Z	CO:Nivel Deposito %	86.067924	88001 Deposito	22880010
2	MARCHA_POZO	1	DEPOSITO 2021-08-01T00:30:00.000Z	CO:Nivel Deposito %	86.240074	88001 Deposito	22880010
3	TOTALIZADOR						
4	PRESION_ENTRADA_SECTOR						
5	VOLUMEN_DIARIO						
6	CAUDAL						
7	PRESION_SALIDA_SECTOR						
8	VOLUMEN						
9	DEPOSITO_METROS						
10	DEPOSITO_VOLUMEN_SALIDA						

2. gis.csv

This file contains information related to the hierarchical and locational relationship of the elements that exist in the water supply service of Villarubia. GIS means Geographic Information System.

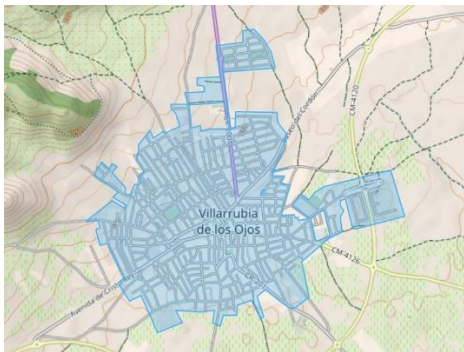
These are the variables in gis.csv dataset:

- **oid**: Identifier of the service and Gis code.
- **bissioCode**: Sector code.
- **Code**: Gis (Geographic Information System) code.
- **parentCode**: code of the parent sector.
- **Name**: Name of the sector.
- **Service**: Service code. This code is always VR (Villarubia).

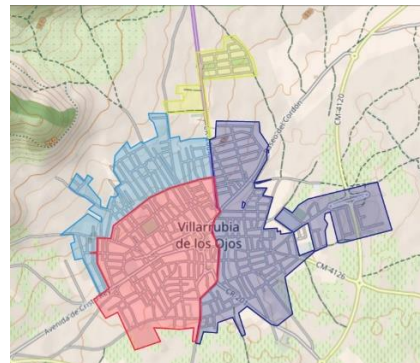
gis.csv example

	oid	bissioCode	code	parentCode	name	service
42	VR22110000	1	22110000	22100000.0	Carretera de Urda	VR
43	VR22120000	2	22120000	22100000.0	Valladolid	VR

Hierarchical Level 1



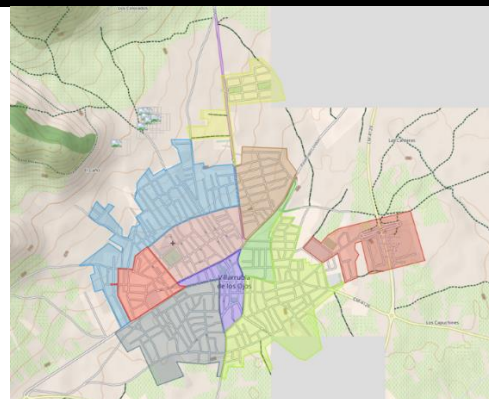
Hierarchical Level 2



Hierarchical Level 3



Hierarchical Level 4



3. abonado.csv

This file contains information related to the provision of service in Villarubia.

These are the variables in abonado.csv dataset:

- **COD_CONCESION**: Service code. This code is always VR (Villarubia).
- **SECTOR**: Sector code.
- **COD_SUMINISTRO**: Contract reference.
- **FECHA_ALTA_SUMINISTRO**: Contract start date.
- **FECHA_BAJA_SUMINISTRO**: Contract termination date.
- **COD_TIPO_SUMINISTRO**: Type of supply.
- **NUM_VIVIENDAS**: Number of dwellings registered in the contract.
- **NUM_PERSONAS**: Number of people registered in the contract.

abonado.csv example								
COD_CONCESION	SECTOR	COD_SUMINISTRO	FECHA_ALTA_SUMINISTRO	FECHA_BAJA_SUMINISTRO	COD_TIPO_SUMINISTRO	NUM_VIVIENDAS	NUM_PERSONAS	
7133	VR	2	1000002	2009-08-13	NaN	CONT	1.0	3.0
7134	VR	0	1000003	2009-08-17	NaN	CONT	1.0	3.0

4. cliente.csv

This file contains information on the type of customer and the model of their meter.

These are the variables in cliente.csv dataset:

- **cod_concesion**: Service code. This code is always VR (Villarubia).
- **cod_tipo_sujeto**: Type of customer.
- **COD_SUMINISTRO**: Contract reference.
- **SECTOR**: Sector code.
- **COD_TIPO_SUMINISTRO**: Type of supply.
- **COD_CONTADOR_ACT**: Meter model.

Cliente.csv example						
cod_concesion	cod_tipo_sujeto	COD_SUMINISTRO	SECTOR	COD_TIPO_SUMINISTRO	COD_CONTADOR_ACT	
7131	VR	PER	1000002	2	CONT	Q18EA019889G
7132	VR	PER	1000003	0	CONT	J14LA719590L

5. contador.csv

This file contains information related to the meter and its manufacturer.

These are the variables in Contador.csv dataset:

- **cod_concesion**: Service code. This code is always VR (Villarubia).
- **COD_CONTADOR_ACT**: model of the meter.
- **CALIBRE**: diameter of the meter.

contador.csv example			
cod_concesion	COD_CONTADOR_ACT	CALIBRE	
7131	VR	Q18EA019889G	13.0
7132	VR	J14LA719590L	13.0

6. clima.csv

This dataset contains information about the weather conditions.

clima.csv example														
	wdir	temp	maxt	visibility	wspd	datetimeStr	solarenergy	heatindex	cloudcover	mint	datetime	precip	solarradiation	weathertype
0	264.50	4.4	7.1	15.3	22.1	2021-01-01T00:00:00+01:00	7.4	NaN	55.3	2.1	1609459200000	2.6	205.2	Drizzle, Mist, Light Drizzle, Rain
1	291.04	2.1	6.9	19.6	18.5	2021-01-02T00:00:00+01:00	8.3	NaN	25.0	-1.3	1609545600000	0.0	229.3	NaN
	snowdepth	sealevelpressure	snow	dew	humidity	precipcover	wgust	conditions	windchill	info	locations			
	0.0		1012.6	0.0	1.8	83.64	12.5	NaN	Rain, Partially cloudy	-1.5	NaN	39.2210171,-3.6185337		
	0.0		1016.4	0.0	-3.3	70.25	0.0	NaN	Clear	-3.9	NaN	39.2210171,-3.6185337		