

EPC data:

Technical documentation

Creation date	2021-09-01
Edition	03
Author	Ellen Webborn, Jessica Few
Project	Smart Energy Research Lab (SERL)
Organisation	University College London (UCL)

Introduction

This document describes the England and Wales Energy Performance Certificate (EPC) data collected for SERL participants, stored in the file *serl_epc_data_edition03.csv*. The data contains 80 columns and 6921 rows (one row per participant with available EPC data). This document lists the EPC variables available along with basic information about the values for each variable such as number of unique values and statistics for numerical variables. A guide to the variables is available [here](#).

Data were collected with the Domestic Energy Performance Certificates API using the house number and postcode (details [here](#)). Where more than one EPC is registered for an address the most recent is provided.

The data have not been modified from the original source except for the removal of address data (replaced with our PUPRN (a unique identifier) used in the other datasets). Note that data quality analysis has not been performed for this dataset.

Data summary

Table 1 lists all variables currently available in the SERL EPC dataset. The number of unique values is given, alongside the R data class and an example value from the dataset.

For variables with fewer than 10 unique values in the EPC dataset, Table 2 shows the number of records with each value and the percent with this value (or non-value in the case of N/A or 'NO DATA!' etc.). We also include PUPRN to show the number of records.

Table 1: All EPC variables, the number of unique values found for each variable, the variable (R) class, and an example from the dataset. In some cases examples are invented (for statistical disclosure control) as the count for unique reads is very low; in these instances examples take the same format as the real data.

<i>variable</i>	<i>n unique values</i>	<i>class</i>	<i>example</i>
PUPRN	6,921	character	1ABC2DE3
currentEnergyRating	7	character	D
potentialEnergyRating	7	character	D
currentEnergyEfficiency	93	integer	60
potentialEnergyEfficiency	99	integer	84
propertyType	5	character	House
builtForm	7	character	Semi-Detached
inspectionDate	3,012	character	13/05/2008
localAuthority	339	character	E07000139
constituency	572	character	E14001013
lodgementDate	3,174	character	28/01/2014
transactionType	17	character	rental (social)
environmentImpactCurrent	97	integer	45
environmentImpactPotential	95	integer	80
energyConsumptionCurrent	586	integer	234
energyConsumptionPotential	500	numeric	77
co2EmissionsCurrent	132	numeric	4.2
co2EmissCurrPerFloorArea	351	numeric	37
co2EmissionsPotential	129	numeric	0.9
lightingCostCurrent	234	numeric	88
lightingCostPotential	156	numeric	33
heatingCostCurrent	1,587	numeric	295
heatingCostPotential	1,209	numeric	356
hotWaterCostCurrent	414	numeric	72
hotWaterCostPotential	262	numeric	57
totalFloorArea	1,702	numeric	80.00
energyTariff	9	character	Single
mainsGasFlag	3	character	Y
floorLevel	25	character	Ground
flatTopStorey	3	character	Y
flatStoreyCount	14	integer	2
mainHeatingControls	47	integer	2107

<i>variable</i>	<i>n unique values</i>	<i>class</i>	<i>example</i>
multiGlazeProportion	83	numeric	100
glazedType	11	character	double glazing, unknown install date
glazedArea	7	character	Much More Than Typical
extensionCount	6	numeric	1
numberHabitableRooms	16	numeric	4
numberHeatedRooms	17	numeric	5
lowEnergyLighting	100	integer	67
numberOpenFireplaces	9	integer	0
hotwaterDescription	38	character	From main system
hotWaterEnergyEff	6	character	Good
hotWaterEnvEff	6	character	Good
floorDescription	93	character	Suspended, no insulation (assumed)
floorEnergyEff	6	character	Very Good
floorEnvEff	5	character	Good
windowsDescription	23	character	Fully double glazed
windowsEnergyEff	6	character	Average
windowsEnvEff	6	character	Poor
wallsDescription	126	character	Cavity wall, as built, no insulation (assumed)
wallsEnergyEff	6	character	Very Good
wallsEnvEff	6	character	Good
secondheatDescription	19	character	None
sheatingEnergyEff	1	character	N/A
sheatingEnvEff	1	character	N/A
roofDescription	119	character	Pitched, no insulation
roofEnergyEff	6	character	Good
roofEnvEff	6	character	N/A
mainheatDescription	51	character	Boiler and radiators, mains gas
mainheatEnergyEff	6	character	Good
mainheatEnvEff	6	character	Good
mainheatcontDescription	44	character	Programmer, room thermostat and TRVs
mainheatcEnergyEff	6	character	Good

<i>variable</i>	<i>n unique values</i>	<i>class</i>	<i>example</i>
mainheatcEnvEff	6	character	Average
lightingDescription	111	character	Low energy lighting in all fixed outlets
lightingEnergyEff	7	character	Good
lightingEnvEff	6	character	Good
mainFuel	31	character	oil (not community)
windTurbineCount	4	numeric	0
heatLossCorridor	5	character	unheated corridor
unheatedCorridorLength	368	numeric	5.189
floorHeight	170	numeric	2.40
photoSupply	16	numeric	0
solarWaterHeatingFlag	3	character	N
mechanicalVentilation	5	character	natural
constructionAgeBand	15	character	England and Wales: 1991-1995
lodgementDatetime	6,887	character	02/10/2010 14:51
tenure	10	character	owner-occupied
fixedLightingOutletsCount	59	numeric	15
lowEnergyFixedLightCount	45	numeric	4

Table 2: The number and percent of each value found in the dataset for each variable with fewer than 10 unique values found. Note that number is rounded down to the nearest 10 for statistical disclosure control, and percent is the rounded number as a percent of total (so may not sum to 100%).

<i>variable</i>	<i>value</i>	<i>number (rounded)</i>	<i>percent</i>
currentEnergyRating	A	0	0.0
	B	440	6.4
	C	2,040	29.5
	D	3,050	44.1
	E	1,130	16.3
	F	180	2.6
	G	50	0.7
potentialEnergyRating	A	180	2.6
	B	3,310	47.8
	C	2,670	38.6

<i>variable</i>	<i>value</i>	<i>number (rounded)</i>	<i>percent</i>
	D	580	8.4
	E	130	1.9
	F	20	0.3
	G	0	0.0
propertyType	Bungalow	850	12.3
	Flat	1,230	17.8
	House	4,690	67.8
	Maisonette	120	1.7
	Park home	0	0.0
builtForm	Detached	1,840	26.6
	Enclosed End-Terrace	80	1.2
	Enclosed Mid-Terrace	50	0.7
	End-Terrace	850	12.3
	Mid-Terrace	1,720	24.9
	NO DATA!	110	1.6
	Semi-Detached	2,250	32.5
energyTariff	INVALID!	0	0.0
	NO DATA!	0	0.0
	Single	5,170	74.7
	Unknown	590	8.5
	dual	630	9.1
	dual (24 hour)	0	0.0
	off-peak 10 hour	10	0.1
	off-peak 7 hour	100	1.4
	standard tariff	380	5.5
mainsGasFlag		400	5.8
	N	730	10.5
	Y	5,770	83.4
flatTopStorey		5,690	82.2
	N	760	11.0
	Y	460	6.6
glazedArea		0	0.0
	Less Than Typical	10	0.1

<i>variable</i>	<i>value</i>	<i>number (rounded)</i>	<i>percent</i>
	More Than Typical	180	2.6
	Much Less Than Typical	0	0.0
	Much More Than Typical	40	0.6
	NO DATA!	400	5.8
	Normal	6,270	90.6
extensionCount		400	5.8
	0	3,970	57.4
	1	1,850	26.7
	2	540	7.8
	3	110	1.6
	4	30	0.4
numberOpenFireplaces		170	2.5
	0	5,880	85.0
	1	670	9.7
	2	140	2.0
	3	20	0.3
	4	10	0.1
	5	0	0.0
	6	0	0.0
	7	0	0.0
hotWaterEnergyEff	Average	950	13.7
	Good	4,730	68.3
	N/A	20	0.3
	Poor	410	5.9
	Very Good	460	6.6
	Very Poor	320	4.6
hotWaterEnvEff	Average	770	11.1
	Good	4,880	70.5
	N/A	20	0.3
	Poor	570	8.2
	Very Good	440	6.4
	Very Poor	220	3.2
floorEnergyEff	Average	0	0.0

<i>variable</i>	<i>value</i>	<i>number (rounded)</i>	<i>percent</i>
	Good	60	0.9
	N/A	3,540	51.1
	NO DATA!	3,090	44.6
	Poor	0	0.0
	Very Good	210	3.0
floorEnvEff	Average	0	0.0
	Good	60	0.9
	N/A	6,630	95.8
	Poor	0	0.0
	Very Good	210	3.0
windowsEnergyEff	Average	3,840	55.5
	Good	2,120	30.6
	N/A	10	0.1
	Poor	320	4.6
	Very Good	260	3.8
	Very Poor	340	4.9
windowsEnvEff	Average	3,840	55.5
	Good	2,120	30.6
	N/A	10	0.1
	Poor	320	4.6
	Very Good	260	3.8
	Very Poor	340	4.9
wallsEnergyEff	Average	580	8.4
	Good	3,170	45.8
	N/A	10	0.1
	Poor	1,120	16.2
	Very Good	330	4.8
	Very Poor	1,690	24.4
wallsEnvEff	Average	580	8.4
	Good	3,170	45.8
	N/A	10	0.1
	Poor	1,120	16.2
	Very Good	330	4.8

<i>variable</i>	<i>value</i>	<i>number (rounded)</i>	<i>percent</i>
	Very Poor	1,690	24.4
sheatingEnergyEff	N/A	6,920	100.0
sheatingEnvEff	N/A	6,920	100.0
roofEnergyEff	Average	1,220	17.6
	Good	2,880	41.6
	N/A	840	12.1
	Poor	340	4.9
	Very Good	700	10.1
	Very Poor	910	13.1
roofEnvEff	Average	1,220	17.6
	Good	2,880	41.6
	N/A	840	12.1
	Poor	340	4.9
	Very Good	700	10.1
	Very Poor	910	13.1
mainheatEnergyEff	Average	570	8.2
	Good	5,610	81.1
	N/A	20	0.3
	Poor	150	2.2
	Very Good	350	5.1
	Very Poor	200	2.9
mainheatEnvEff	Average	300	4.3
	Good	5,680	82.1
	N/A	20	0.3
	Poor	160	2.3
	Very Good	430	6.2
	Very Poor	310	4.5
mainheatcEnergyEff	Average	2,210	31.9
	Good	3,630	52.4
	N/A	20	0.3
	Poor	460	6.6
	Very Good	200	2.9
	Very Poor	390	5.6

<i>variable</i>	<i>value</i>	<i>number (rounded)</i>	<i>percent</i>
mainheatcEnvEff	Average	2,210	31.9
	Good	3,630	52.4
	N/A	20	0.3
	Poor	460	6.6
	Very Good	200	2.9
	Very Poor	390	5.6
lightingEnergyEff		0	0.0
	Average	1,180	17.0
	Good	1,390	20.1
	N/A	10	0.1
	Poor	720	10.4
	Very Good	2,560	37.0
	Very Poor	1,050	15.2
lightingEnvEff	Average	1,180	17.0
	Good	1,390	20.1
	N/A	10	0.1
	Poor	720	10.4
	Very Good	2,560	37.0
	Very Poor	1,050	15.2
windTurbineCount		220	3.2
	-1	0	0.0
	0	6,680	96.5
	1	0	0.0
heatLossCorridor		4,500	65.0
	NO DATA!	2,020	29.2
	heated corridor	60	0.9
	no corridor	140	2.0
	unheated corridor	180	2.6
solarWaterHeatingFlag		2,610	37.7
	N	4,270	61.7
	Y	30	0.4
mechanicalVentilation		0	0.0
	NO DATA!	400	5.8

<i>variable</i>	<i>value</i>	<i>number (rounded)</i>	<i>percent</i>
	mechanical, extract only	30	0.4
	mechanical, supply and extract	10	0.1
	natural	6,470	93.5