

Supplement of Adv. Geosci., 45, 377–382, 2018
<https://doi.org/10.5194/adgeo-45-377-2018-supplement>
© Author(s) 2018. This work is distributed under
the Creative Commons Attribution 4.0 License.



Supplement of

Accuracy measurement of Random Forests and Linear Regression for mass appraisal models that estimate the prices of residential apartments in Nicosia, Cyprus

Thomas Dimopoulos et al.

Correspondence to: Thomas Dimopoulos (thomas.dimopoulos@gmail.com)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

Figure S1: Distribution of NDEV per Municipality

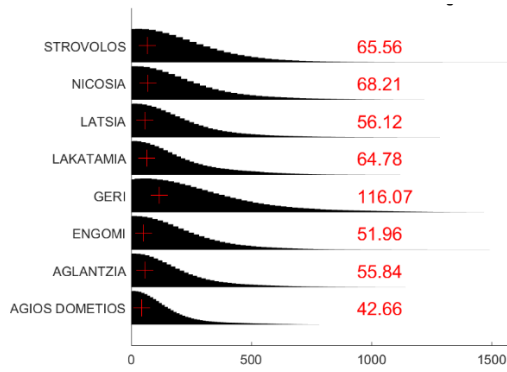


Figure S2: Relief weights for all the values of DEV

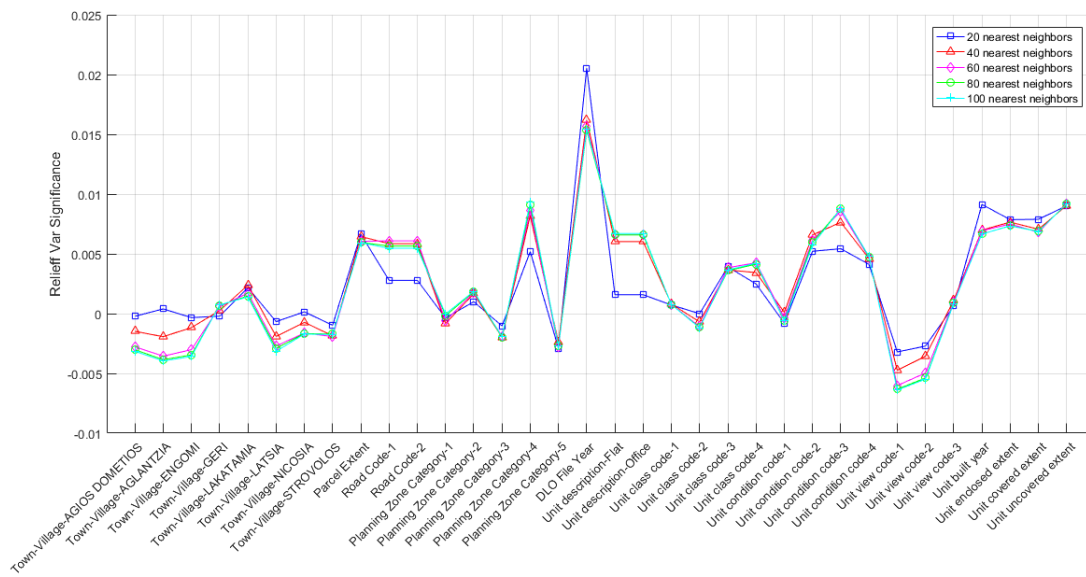


Figure S3: Relief weights for the non-zero values of DEV

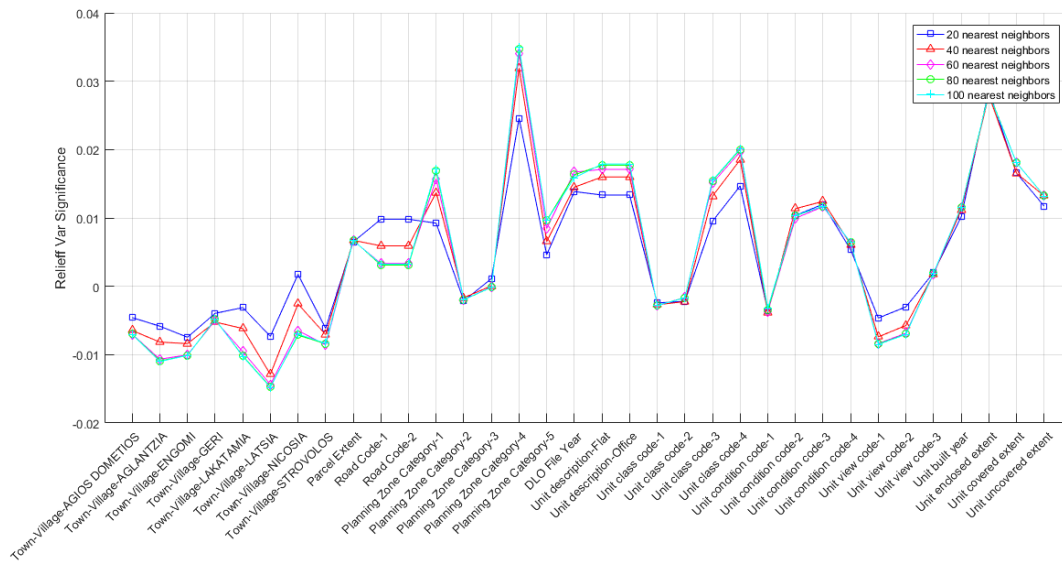


Table S1: Formulas utilized in the models

Number	Investigated Formula
1	Accepted_price ~ Town_village_name_mod
2	Accepted_price ~ Road_code
3	Accepted_price ~ Shape_code
4	Accepted_price ~ Unit_class_code
5	Accepted_price ~ Unit_condition_code
6	Accepted_price ~ Unit_condition_mod_code
7	Accepted_price ~ Unit_view_code
8	Accepted_price ~ Unit_enclosed_extent
9	Accepted_price ~ Unit_covered_extent
10	Accepted_price ~ Unit_uncovered_extent
11	Accepted_price ~ Unit_encl_and_covered_extent
12	Accepted_price ~ Unit_total_extent
13	Accepted_price ~ Unit_adjusted_extent
14	Accepted_price ~ price_index_sale_acceptance_date
15	Accepted_price ~ flats_index_sale_acceptance_date
16	Accepted_price ~ Sale_acceptance_year
17	Accepted_price ~ age_at_sale
18	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent
19	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + Unit_uncovered_extent
20	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + Town_village_name_mod
21	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + Unit_class_code
22	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + Unit_condition_code
23	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + Unit_condition_mod_code
24	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + price_index_sale_acceptance_date
25	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + flats_index_sale_acceptance_date
26	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + Sale_acceptance_year
27	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale
28	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Town_village_name_mod
29	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_class_code
30	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_code
31	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code

32	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + price_index_sale_acceptance_date
33	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + flats_index_sale_acceptance_date
34	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Sale_acceptance_year
35	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + Town_village_name_mod
36	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + Unit_class_code
37	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + price_index_sale_acceptance_date
38	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + flats_index_sale_acceptance_date
39	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + Sale_acceptance_year
40	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + flats_index_sale_acceptance_date + Town_village_name_mod
41	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + flats_index_sale_acceptance_date + Unit_class_code
42	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + flats_index_sale_acceptance_date + Unit_class_code + Town_village_name_mod
43	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + flats_index_sale_acceptance_date + Unit_class_code + Town_village_name_mod + Unit_uncovered_extent
44	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + flats_index_sale_acceptance_date + Unit_class_code + Town_village_name_mod + Unit_uncovered_extent + price_index_sale_acceptance_date
45	Accepted_price ~ Unit_enclosed_extent + Unit_covered_extent + age_at_sale + Unit_condition_mod_code + flats_index_sale_acceptance_date + Unit_class_code + Town_village_name_mod + Unit_uncovered_extent + price_index_sale_acceptance_date + Unit_view_code

Table S2: Accuracy Measures for Linear Regression

model	lin_coef	RMSE	MAE	MAPE	sr	method
1	0.05	53216.5	42414.9	0.39	1.23	Linear Regression
2	0	55032	44374.4	0.4	1.24	Linear Regression
3	0	54683.4	44076.5	0.4	1.23	Linear Regression
4	0.04	52935.8	41734.4	0.38	1.21	Linear Regression
5	0.12	51147.5	40382.8	0.37	1.21	Linear Regression
6	0.12	51213	40508.9	0.37	1.21	Linear Regression
7	0.05	52198.7	42011.9	0.38	1.21	Linear Regression

8	0.45	43077.7	33357	0.27	1.12	Linear Regression
9	0.33	45443.4	34504.8	0.3	1.16	Linear Regression
10	0.07	52615.6	42011.3	0.38	1.22	Linear Regression
11	0.53	40859.7	30740.6	0.24	1.1	Linear Regression
12	0.51	41193.3	31209.6	0.25	1.12	Linear Regression
13	0.52	41146.1	31448.4	0.25	1.11	Linear Regression
14	0	54683	44027.9	0.4	1.24	Linear Regression
15	0	54793.9	44097.9	0.41	1.24	Linear Regression
16	0	55198.7	44376.3	0.41	1.24	Linear Regression
17	0.15	49313	39069.6	0.35	1.19	Linear Regression
18	0.57	39678.8	28811.1	0.23	1.1	Linear Regression
19	0.6	38436.2	28169.8	0.22	1.1	Linear Regression
20	0.65	37996.5	28112	0.22	1.08	Linear Regression
21	0.55	38447.9	28313.1	0.22	1.09	Linear Regression
22	0.65	35698.7	26481.4	0.2	1.08	Linear Regression
23	0.65	35799.8	26539.2	0.21	1.08	Linear Regression
24	0.59	39641.8	29273.4	0.23	1.09	Linear Regression
25	0.58	39875	29248.4	0.23	1.1	Linear Regression
26	0.59	39614.4	29373.6	0.23	1.1	Linear Regression
27	0.73	31621.4	23221.8	0.18	1.06	Linear Regression
28	0.74	31232.6	22887.1	0.18	1.06	Linear Regression
29	0.73	31702	23502.1	0.19	1.06	Linear Regression
30	0.76	31519	23399.4	0.19	1.05	Linear Regression
31	0.75	31774	23472.9	0.19	1.06	Linear Regression
32	0.74	31920.5	23607.3	0.19	1.06	Linear Regression
33	0.73	31708.8	23312.6	0.18	1.06	Linear Regression
34	0.74	31603.9	23266.5	0.18	1.05	Linear Regression
35	0.76	31547.7	23209.7	0.18	1.06	Linear Regression
36	0.75	31927.8	23676.7	0.19	1.06	Linear Regression
37	0.75	31974.5	23742.5	0.19	1.06	Linear Regression
38	0.75	31767.5	23465.3	0.19	1.06	Linear Regression
39	0.75	31776.2	23583.4	0.18	1.05	Linear Regression
40	0.76	31512.1	23176.3	0.18	1.06	Linear Regression
41	0.75	31918.3	23665.7	0.19	1.06	Linear Regression
42	0.76	31667.2	23411.6	0.19	1.06	Linear Regression
43	0.77	30868.9	23027.4	0.18	1.06	Linear Regression
44	0.78	30732	23106.5	0.18	1.05	Linear Regression
45	0.79	30495.7	22973.8	0.18	1.05	Linear Regression

Table S3: Accuracy Measures for the Random Forests Method

model	lin_coef	RMSE	MAE	MAPE	sr	method
1	0.05	53211.2	42398.7	0.39	1.23	Random Forests

2	0	55001.3	44342.1	0.4	1.24	Random Forests
3	0	54648.9	44019.9	0.4	1.23	Random Forests
4	0.04	52939.6	41736.9	0.38	1.21	Random Forests
5	0.12	51135.3	40365.5	0.36	1.21	Random Forests
6	0.12	51232.2	40543.9	0.37	1.21	Random Forests
7	0.05	52202.5	42018.6	0.38	1.21	Random Forests
8	0.46	46422.5	34488.6	0.27	1.11	Random Forests
9	0.41	48011.5	36344.7	0.31	1.16	Random Forests
10	0.09	54920.8	43351.6	0.38	1.2	Random Forests
11	0.5	45033	31782.6	0.24	1.09	Random Forests
12	0.59	45231.8	31994.7	0.25	1.1	Random Forests
13	0.5	44896.3	33133.9	0.26	1.1	Random Forests
14	0.01	59853.3	47559.5	0.43	1.24	Random Forests
15	0.01	56177.9	45280	0.42	1.25	Random Forests
16	0	55192.8	44356.3	0.41	1.24	Random Forests
17	0.23	62465.3	44821	0.36	1.15	Random Forests
18	0.6	44138.9	29343	0.22	1.09	Random Forests
19	0.58	38781.9	27819.4	0.22	1.1	Random Forests
20	0.56	40920.1	27676.9	0.22	1.1	Random Forests
21	0.48	37838.1	28108.4	0.23	1.11	Random Forests
22	0.54	37359.4	26894.2	0.22	1.1	Random Forests
23	0.52	37173	27065.2	0.22	1.11	Random Forests
24	0.55	41474.8	28768.3	0.22	1.09	Random Forests
25	0.55	40426.3	28831	0.22	1.09	Random Forests
26	0.52	41408.3	29358.4	0.23	1.1	Random Forests
27	0.69	30659.6	21574.3	0.17	1.06	Random Forests
28	0.61	31166.7	21881.3	0.17	1.07	Random Forests
29	0.57	31182.6	23234	0.19	1.09	Random Forests
30	0.57	32217.1	23674.4	0.19	1.09	Random Forests
31	0.57	32244.9	23779	0.19	1.09	Random Forests
32	0.65	30730.7	21856	0.17	1.07	Random Forests
33	0.66	29845.2	21583.6	0.17	1.07	Random Forests
34	0.62	30843.9	22166.5	0.18	1.08	Random Forests
35	0.55	32365.3	23636.7	0.19	1.09	Random Forests
36	0.52	32518	24520.5	0.2	1.1	Random Forests
37	0.55	32605.8	24089.8	0.2	1.1	Random Forests
38	0.57	31605.6	23487.6	0.19	1.09	Random Forests
39	0.54	32362.1	24101.8	0.2	1.1	Random Forests
40	0.71	29418.4	20132.1	0.15	1.06	Random Forests
41	0.7	29499.5	20647.8	0.16	1.06	Random Forests
42	0.69	29050	20133.3	0.15	1.06	Random Forests
43	0.7	27603.7	19760.1	0.15	1.06	Random Forests

44	0.71	28122.1	19856.6	0.15	1.06	Random Forests
45	0.71	27755.9	19723.8	0.15	1.07	Random Forests

Table S4: Accuracy Measures Comparison

model	lin_coef	RMSE	MAE	MAPE	sr	method
1	0.000	0.010	0.038	0.000	0.000	comparison
2	-	0.056	0.073	0.000	0.000	comparison
3	-	0.063	0.128	0.000	0.000	comparison
4	0.000	-0.007	-0.006	0.000	0.000	comparison
5	0.000	0.024	0.043	2.740	0.000	comparison
6	0.000	-0.037	-0.086	0.000	0.000	comparison
7	0.000	-0.007	-0.016	0.000	0.000	comparison
8	-2.198	-7.474	-3.336	0.000	0.897	comparison
9	-21.622	-5.496	-5.194	-3.279	0.000	comparison
10	-	-4.287	-3.140	0.000	1.653	comparison
11	5.825	-9.717	-3.333	0.000	0.913	comparison
12	-14.545	-9.346	-2.484	0.000	1.802	comparison
13	3.922	-8.717	-5.220	-3.922	0.905	comparison
14	-	-9.028	-7.712	-7.229	0.000	comparison
15	-	-2.494	-2.645	-2.410	-0.803	comparison
16	-	0.011	0.045	0.000	0.000	comparison
17	-42.105	-23.533	-13.712	-2.817	3.419	comparison
18	-5.128	-10.642	-1.829	4.444	0.913	comparison
19	3.390	-0.895	1.251	0.000	0.000	comparison
20	14.876	-7.409	1.560	0.000	-1.835	comparison
21	13.592	1.599	0.725	-4.444	-1.818	comparison
22	18.487	-4.546	-1.547	-9.524	-1.835	comparison
23	22.222	-3.764	-1.963	-4.651	-2.740	comparison
24	7.018	-4.519	1.740	4.444	0.000	comparison
25	5.310	-1.373	1.437	4.444	0.913	comparison
26	12.613	-4.428	0.052	0.000	0.000	comparison
27	5.634	3.088	7.355	5.714	0.000	comparison
28	19.259	0.211	4.493	5.714	-0.939	comparison
29	24.615	1.652	1.147	0.000	-2.791	comparison
30	28.571	-2.191	-1.169	0.000	-3.738	comparison
31	27.273	-1.471	-1.296	0.000	-2.791	comparison
32	12.950	3.798	7.704	11.111	-0.939	comparison
33	10.072	6.055	7.702	5.714	-0.939	comparison
34	17.647	2.434	4.842	0.000	-2.817	comparison
35	32.061	-2.559	-1.823	-5.405	-2.791	comparison
36	36.220	-1.832	-3.501	-5.128	-3.704	comparison
37	30.769	-1.955	-1.452	-5.128	-3.704	comparison

38	27.273	0.511	-0.095	0.000	-2.791	comparison
39	32.558	-1.827	-2.174	-10.526	-4.651	comparison
40	6.803	6.873	14.058	18.182	0.000	comparison
41	6.897	7.876	13.621	17.143	0.000	comparison
42	9.655	8.621	15.057	23.529	0.000	comparison
43	9.524	11.168	15.272	18.182	0.000	comparison
44	9.396	8.869	15.129	18.182	-0.948	comparison
45	10.667	9.407	15.224	18.182	-1.887	comparison
%diff	9.730765	-1.27179	1.443673	2.072518	-0.73431	

Table S5: Description of the fields in the studied Database

A	Sbpi_id_no	The unique identification number for the CILIS database (GIS).	Integer
B	District_name	The district name.	String
C	Town_village_name	It identifies the municipality / village.	String
D	Quarter_name	Separation of municipalities into quarters. The separation is based on the closest church (Parish)	String
E	District	The district name of the variable District_name is coded numerically.	Integer
F	Town_village	The municipality/village name of the variable Town_village_name is coded numerically.	Integer
G	Quarter	The quarter name of the variable Quarter_name is coded numerically.	Integer
H	Block	The plans are divided into plots.	Integer
I	Reg_no	The Unique registration number.	Integer
J	Sheet	Cyprus cadastral map is divided in 60 Sheets (scale 1:5000, 2 inches = 1 mile).	Integer
K	Plan	Every sheet is further divided into 4 plans or smaller (scales 1:2500, 1:1000, 1:500).	String
L	Parcel_no	The parcel number. Parcel (or plot) is the minimum division.	Integer
M	Parcel_extent	Area of the parcel (or plot).	Double
N	Access_code	It denotes the access to the parcel. Values: 1 = access through public road. Since all data are referring to developed land, all of them should be accessible.	Integer

O	Road_code	The importance of the adjacent roads. Values: From 1 (most important) to 2 (less important).	Integer
P	Shape_code	It denotes the quality of the shape of the parcel. Values: From 1 (best shape) to 3 (worst shape).	Integer
Q	Planning_zone_category	Values: 1 = residential, 2 = mixed use (offices and residential), 3 = core (continuous density old city centres), 4 = commercial and bussines district, 5 = commercial, 8 = industrial.	Integer
R	Planning_zone_code	The code of the planning zone.	Integer
S	Planning_zone_name	The name of the planning zone. It corresponds to the Planning_zone_code variable.	String
T	Secondary_planning_zone_category	A secondary category for the planning zone, in case the parcel abuts in 2 different zones. Values: 1 = residential, 2 = mixed use (offices and residential), 3 = core (continuous density old city centres), 4 = commercial and bussines district, 5 = commercial.	Integer
U	Secondary_planning_zone_code	The code of the secondary planning zone in case the parcel abuts in 2 different zones.	Integer
V	Secondary_planning_zone_name	The name of the secondary planning zone. It corresponds to the Secondary_planning_zone_code variable.	String
W	Declared_price	The corresponding amount (in €) that the buyer and seller declared at the Cyprus Department of Lands and Surveys (DLS) for the property exchange at the transaction date. No changes made for inflation, property cycles and/or fluctuation.	Double

X	Accepted_price	The desktop valuation (in €) of DLS (transaction fees and taxes are paid on this figure). Theoretically, this figure must be more accurate than the Declared_price. However due to the fact that the DLS is doing retrospective research, when property prices are increasing they are not up to date (they evaluate a bit lower than real prices). On the contrary when property market is not performing well, when they look back they have higher figures than the real ones. No changes made for inflation, property cycles and/or fluctuation.	Double
Y	DLO_file	All transactions are marked with Greek Letter "Π" and the sale number (see Z column).	String
Z	DLO_file_no	The number of the sale (like a receipt).	Integer
AA	DLO_file_year	The year the transaction happened.	Integer
AB	Fiscal_property_type	Horizontal (applicable mainly to apartment buildings). It applies when there is ownership of separate parts of a multi-storey building.	String
AC	Sale_acceptance_date	The date that transaction was accepted from DLS.	Date
AD	Share_numerator	Since the sample under study consists of flats, the majority of observations are referring to 100% shares. Those observations that are not 100% share are excluded from the sample because a 50% share of a flat does not represent a rational market value.	Integer
AE	Share_denominator	In case of share. See also column AD	Integer
AF	COS_agreement_date	The date that the involved parties submitted the contract to the DLS.	Date
AG	Remark	DLS notes.	String

AH	Sales_remark	DLS notes in Greek. In some cases, they repeat the declared price or the covered area.	String
AI	Main_sbp_cat	Notes of the DLS.	String
AJ	Main_sbp_kind	The type of property.	String
AK	Status	All observations are marked as L.	String
AL	Main_sbpi_id_no	A code from the DLS system corresponding to the shapefile.	Integer
AM	Building_code	DLS internal reference.	Integer
AN	Building_desc	DLS internal reference.	String
AO	Building_sbpi_id_no	DLS internal reference.	Integer
AP	Field42	DLS internal reference.	String
AQ	Unit_code	The unit code that corresponds to the Unit_desc variable in column AR. Values: 215 = apartment, 818 = coffee shop, 814 = exposition centre, 823 = health clinic, 815 = laboratory, 810 = office, 811 = office complex, 813 = office with middle floor, 837 = power station, 853 = residence, 474 = shed, 808 = shop, 812 = shop with middle floor, 809 = shopping complex, 927 = small industry, 849 = storage, 932 = three storey residence, 824 = treatment room, 232 = two storey apartment, 861 = two storey building, 803 = two storey residence, 805 = two storey residence with garden, 237 = unfinished building.	Integer
AR	Unit_desc	The type of property. Values: apartment, coffee shop, exposition center, health clinic, laboratory, office, office complex, office with middle floor, power station, residence, shed, shop, shop with middle floor, shopping complex, small industry, storage, three storey residence, treatment room, two storey apartment, two storey building, two storey residence, two storey residence with garden, unfinished building.	String
AS	Unit_sbpi_id_no	A code from the DLS system corresponding to the shapefile.	Integer

AT	Field46	Notes of the DLS.	String
AU	Unit_class_code	It denotes the class of the building. Values: From 1 (best class) to 4 (worst class).	Integer
AV	Unit_condition_code	It denotes the condition of the building. Values: From 1 (best condition) to 4 (worst condition).	Integer
AW	Unit_view_code	It denotes the view of the unit. Values: From 1 (best view) to 4 (worst view).	Integer
AX	Unit_built_year	The date the building was constructed.	Date
AY	Unit_enclosed_extent	Internal area (in m2).	Integer
AZ	Unit_covered_extent	Area of covered verandahs (in m2).	Integer
BA	Unit_uncovered_extent	Area of uncovered verandahs (in m2).	Integer
BB	Total_value_2013	The new general valuation (in €) dated the 1st of January 2013. These values represent the outcome of the new general valuation made by the DLS.	Integer