**Roman Fedoseev** 

# Cost of new residential property

### Introduction

The subject of my project is to analyze the cost of new residential property in Luxembourg. There appears to be a strong and widely recognized dependency of the prices to the distance from Luxembourg City. This dependency is the main motivation of this study which attempts to identify other factors and explanatory variables and to propose a multiple linear regression model of the property prices.

## Data

The data are taken from a periodic publication of the prices of the new residential properties available for sale and grouped by region (Center, North, South, West and East) and includes:

- Name of the location
- Type of property (apartment or house)
- Number of bedrooms
- Area of the property (in m<sup>2</sup>)
- Price of the property (in Euros)

The distances from a given location to Luxembourg City are published by the government.<sup>1</sup>

In total, prices for 425 different properties were published in a given publication in 2011, grouped by region as detailed in the table below. 87 properties (20% of the total) are located in Luxembourg City.

Table	1	Data	Cou	nt	

Region	Incl. L-City	Excl. L-City
Center	154	67
South	79	79
North	109	109
West	32	32
East	51	51
Total	425	338

Full data excluding Luxembourg City is listed in Appendix.

## Analysis

As the prices vary greatly between small and large properties, the common practice is to measure the prices in Euros per square meter. We define the response variable  $\mathbf{Y}$  as the prices in Euros per square meter. We start with the first explanatory variable  $\mathbf{X}_1$  for the distance from Luxembourg City, in km.

As seen from Chart 1, the property prices in Luxembourg City (at 0 km distance) are considerably higher than in the rest of the country and they have larger variation. This may be explained by a larger heterogeneity of various factors affecting the price, such as: higher variation of the price of the land within the city, city district, infrastructure, higher variation in cost of construction itself, higher demand (in particular for higher class property), etc. Based on these considerations, the Luxembourg City is excluded from further analysis.

<sup>&</sup>lt;sup>1</sup> The government lists the distances between all pairs of communities in the country and uses these to determine tax deductions for the cost of commuting between place of residence and work. The distances are calculated along a straight line (bird's flight) between the centers of communities.

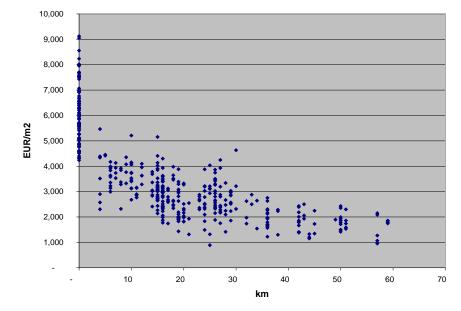


Chart 1 Prices in EUR/m2 as a function of distance from Luxembourg City (all data)

The price of the property has two components: the price of the land and of the construction itself. While one can expect that the land becomes cheaper as the distance from the capital city increases (but still would have a certain limit), the prices of the construction should rather remain consistent throughout the country. Therefore, one would not expect a linear relation between prices per meter squared and the distance, and one can indeed discern a convexity in the chart.

According to Tukey and Mosteller's bulging rule, this particular non-linearity can be corrected by power transformation of either **X** or **Y** or both in the direction of decrease of the power. The Chart 2 shows the plot of ln(Y) as a function of  $X^{1/2}$ :

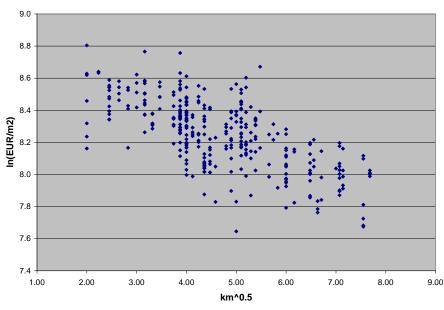


Chart 2 Power transformation Y to lnY and X to  $\sqrt{X}$ 

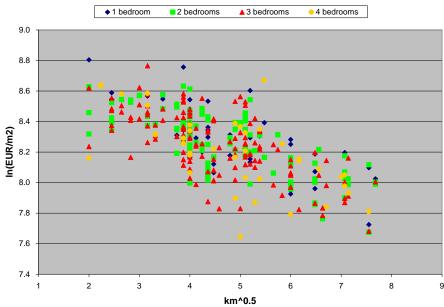
For simplicity of further analysis, the designations  $X_1$  and Y are kept for the transformed variables.

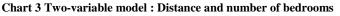
The following table provides the summary of the analysis of variance (ANOVA) for this initial single-variable model.

Regression Si	tatistics					
Multiple R	0.67353					
R Square	0.45365					
Adjusted R Square	0.45202					
Standard Error	0.15841					
Observations	338					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	1	7.00091	7.00091	278.98743	5.05853E-46	
Residual	336	8.43159	0.02509			
Total	337	15.43250				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	8.76120	0.03160	277.25710	0	8.69904	8.82336
X Variable 1	-0.10942	0.00655	-16.70292	5.05853E-46	-0.12231	-0.09654

The R<sup>2</sup> of the regression is equal to 45.365%, and the *t* statistic of -16.7 (or the F-test of 279) with the p-value of  $5 \times 10^{-46}$  show that the null hypothesis (the slope B=0) can be rejected. Therefore, this simple regression is highly significant and confirms the commonly recognized decrease of prices as the distance from Luxembourg City increases (the slope B is negative).

Nevertheless, there is significant variation not explained by the distance alone. Based on the available data, it would be possible to consider the number of bedrooms as an additional explanatory variable  $X_2$ , as shown in the chart below.





The correlation between the two variables (distance and number of bedrooms) is only -3%, which means that variables are highly independent.

The table below shows the ANOVA results of the two-variable model.

#### Table 3 ANOVA for the two-variable model

Regression Statistics					
Multiple R	0.68300				
R Square	0.46649				
Adjusted R Square	0.46330				
Standard Error	0.15677				
Observations	338				

ANOVA

	df	SS	MS	F	Significance F
Regression	2	7.19910	3.59955	146.45826	1.97895E-46
Residual	335	8.23340	0.02458		
Total	337	15.43250			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	8.83120	0.03982	221.77900	0	8.75287	8.90953
X Variable 1	-0.10994	0.00649	-16.95020	5.6804E-47	-0.12270	-0.09718
X Variable 2	-0.02708	0.00954	-2.83969	0.00479	-0.04584	-0.00832

As expected, the  $R^2$  increased (from 45.37% to 46.65%) with the addition of the second explanatory variable. The p-value of the variable  $X_2$  is much smaller than 1, and the slope coefficient is non-zero with at least 95% confidence. Therefore, the number of bedrooms variable appears to be significant for the model.

As the data shows a strong correlation between the number of bedrooms and the type of property (apartment or house), with the correlation coefficient of 57%, it is interesting to study the model where the dummy variable  $\mathbf{D}_1$  for the type of property (1 for house, 0 for apartment) is used instead of number of bedrooms. The ANOVA summary for this model is given in the table below:

#### Table 4 ANOVA for the model with one quantitative and one dummy variable.

Regression S	Statistics					
Multiple R	0.67356					
R Square	0.45368					
Adjusted R Square	0.45042					
Standard Error	0.15864					
Observations	338					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	2	7.00143	3.50072	139.09757	1.0525E-44	
Residual	335	8.43106	0.02517			
Total	337	15.43250				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	8.76208	0.03223	271.86373	0	8.69868	8.82548
X1	-0.10943	0.00656	-16.67918	6.783E-46	-0.12234	-0.09653
D1	-0.00268	0.01860	-0.14415	0.88547	-0.03927	0.03391

One can see that the addition of the dummy variable does not improve the  $R^2$  of the regression and that this variable is in fact insignificant: its p-value is 0.885, and the confidence interval of the slope value includes zero (-0.039; 0.034).

The same conclusion about the insignificance of the type of property dummy variable can be drawn from the analysis of the model with both  $X_2$  and  $D_1$ . Indeed, as seen from the table below, the p-value of  $D_1$  is 0.071, and the null hypothesis is not rejected based on the 95% confidence interval, since it includes the value of zero.

Table 5 ANOVA for the model with two quantitative variables and one dummy variable.

	df	SS	MS	F	Significance F	
Regression	3	7.27916	2.42639	99.39648	5.33608E-46	
Residual	334	8.15334	0.02441			
Total	337	15.43250				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
	0000000000000	olandara Enoi				
Intercept	8.84898		216.45563	0	8.76856	8.92939
Intercept X1				0 3.17539E-47	8.76856 -0.12275	
•	8.84898	0.04088		•		8.92939

ANOVA

Therefore, the dummy variable  $D_1$  for the type of property is not selected for the model.

Finally, it would be interesting to study the differences between regions. First, a set of dummy variables as shown in the table below can be introduced.

#### Table 6 Dummy variables for the region

Region	D2	D3	D4	D5
Center	0	0	0	0
North	1	0	0	0
South	0	1	0	0
West	0	0	1	0
East	0	0	0	1

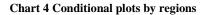
In this model, we are effectively only varying the intercept, while keeping the slope of each regression line (plane) the same.

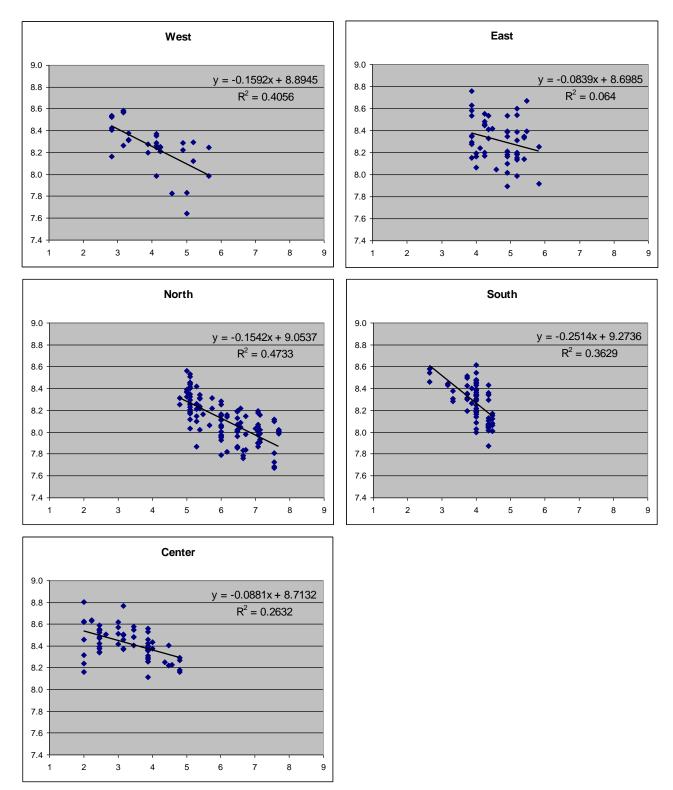
The following ANOVA summary is obtained for the model with two quantitative variables  $(X_1, X_2)$  and four dummy variables  $(D_2, D_3, D_4, D_5)$ :

### Table 7 ANOVA for the model with 2 quantitative variables and 4 dummy variables

Regression S	Statistics					
Multiple R	0.72397					
R Square	0.52413					
Adjusted R Square	0.51550					
Standard Error	0.14895					
Observations	338					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	6	8.08862	1.34810	60.76107	1.63118E-50	
Residual	331	7.34388	0.02219			
Total	337	15.43250				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	8.97864	0.04838	185.60361	0	8.88347	9.07380
X1	-0.13968	0.01104	-12.65190	3.38585E-30	-0.16139	-0.11796
X2	-0.03542	0.00932	-3.80108	0.00017	-0.05374	-0.01709
D2	0.07514	0.03874	1.93946	0.05330	-0.00107	0.15136
D3	-0.07528	0.02639	-2.85281	0.00461	-0.12719	-0.02337
D4	-0.06487	0.03296	-1.96843	0.04985	-0.12970	-4.18455E-05
D5	0.07223	0.03160	2.28604	0.02288	0.01008	0.13438

Although one can note the improvement of  $\mathbb{R}^2$  to 52.41%, the significance of dummy variable  $\mathbb{D}_2$  can be questioned, given the p-value of 0.0533. Furthermore, this model assumes the constant slope coefficient across all regions and only varies the intercept point. However, as seen below, the analysis of the price-distance plots by regions suggests that different regions have different slope coefficients.





It would therefore seem reasonable to introduce interactions between the distance variable  $(X_1)$  and each of the dummy variables  $(D_2 \text{ through } D_5)$ . The results of regression of the model with two quantitative variables  $(X_1, X_2)$ , four dummy variables  $(D_2, D_3, D_4, D_5)$  and four interactions  $(X_1D_2, X_1D_3, X_1D_4, X_1D_5)$  are given in the following table:

### Table 8 ANOVA for the model with 2 quantitative variables, 4 dummy variables and 4 interactions

Regression Statistics				
Multiple R	0.74484			
R Square	0.55478			
Adjusted R Square	0.54117			
Standard Error	0.14495			
Observations	338			

#### ANOVA

	df	SS	MS	F	Significance F
Regression	10	8.56168	0.85617	40.74721	1.06362E-51
Residual	327	6.87082	0.02101		
Total	337	15.43250			

ļ	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	
Intercept	8.80674	0.07582	116.16059	0.00000	8.65760	8.95589	
X1	-0.08253	0.02152	-3.83477	0.00015	-0.12488	-0.04019	
X2	-0.04144	0.00916	-4.52421	0.00001	-0.05946	-0.02342	
D2	0.37406	0.12280	3.04618	0.00251	0.13249	0.61562	
D3	0.64206	0.17843	3.59834	0.00037	0.29104	0.99309	
D4	0.20549	0.14217	1.44541	0.14930	-0.07419	0.48517	
D5	-0.02701	0.17824	-0.15151	0.87966	-0.37765	0.32364	
X1D2	-0.07547	0.02685	-2.81073	0.00524	-0.12829	-0.02265	
X1D3	-0.19037	0.04605	-4.13400	0.00005	-0.28096	-0.09978	
X1D4	-0.07761	0.03676	-2.11144	0.03549	-0.14993	-0.00530	
X1D5	0.00422	0.04067	0.10381	0.91738	-0.07579	0.08423	

The  $R^2$  of this model is 55.48%, and the significance is high (F = 40.74). However, as seen in the plots above, the East data exhibits considerable variation, and has a slope and intercept comparable to those of the Center region. The ANOVA results are in line with these observations, showing high p-values and small slope coefficients for  $D_5$  and  $X_1D_5$ . Based on these considerations, the model is refined by combining East with Center, i.e. removing the  $D_5$  and  $X_1D_5$  variables. The regression results of this model are as follows:

Table 9 ANOVA for the model with 2 quantitative variables, 3 dummy variables and 3 int	eractions
--	-----------

D20.370000.111903.306610.001050.149870.59012D30.637810.170683.737010.000220.302060.97357D40.201470.132791.517170.13018-0.059760.46270X1D2-0.073620.02077-3.544490.00045-0.11447-0.03276X1D3-0.188450.04261-4.422790.00001-0.27227-0.10463							
R Square   0.55469     Adjusted R Square   0.54386     Standard Error   0.14453     Observations   338     ANOVA   F   Significance F     Regression   8   8.56023   1.07003   51.22610   2.13679E-53     Residual   329   6.87227   0.02089   2.13679E-53     Total   337   15.43250   15.61465   0.00000   8.69599   8.92462     X1   -0.08437   0.01330   -6.34591   0.00000   -0.11052   -0.05822     X2   -0.04127   0.00911   -4.53210   0.00001   -0.05919   -0.0236     D2   0.37000   0.11190   3.30661   0.01015   0.14987   0.59012     D3   0.63781   0.17068   3.73701   0.00022   0.30206   0.97357     D4   0.20147   0.13279   1.51717   0.13018   -0.05976   0.46270     X1D2   -0.07362   0.02077   -3.54449   0.00045   -0.11447   -0.03276     X1D3   -0.18845   0.04261   -4.42279   0.00001   -	Regression S	Statistics					
Adjusted R Square     0.54386       Standard Error     0.14453       Observations     338       ANOVA     Image: Construction of the state	Multiple R	0.74477					
Standard Error     0.14453       Observations     338       ANOVA       df     SS     MS     F     Significance F       Regression     8     8.56023     1.07003     51.22610     2.13679E-53       Residual     329     6.87227     0.02089     0.14453       Total     337     15.43250     Upper 95%       Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.011052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.000045 <th< td=""><td>R Square</td><td>0.55469</td><td></td><td></td><td></td><td></td><td></td></th<>	R Square	0.55469					
Observations     338       ANOVA     df     SS     MS     F     Significance F       Regression     8     8.56023     1.07003     51.22610     2.13679E-53       Residual     329     6.87227     0.02089     2.13679E-53       Total     337     15.43250     Upper 95%     Upper 95%       Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.00045     -0.11447     -0.03276	Adjusted R Square	0.54386					
ANOVA       df     SS     MS     F     Significance F       Regression     8     8.56023     1.07003     51.22610     2.13679E-53       Residual     329     6.87227     0.02089     2.13679E-53       Total     337     15.43250     15.43250       Coefficients     Standard Error     t Stat     P-value     Lower 95%     Upper 95%       Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077 <td< td=""><td>Standard Error</td><td>0.14453</td><td></td><td></td><td></td><td></td><td></td></td<>	Standard Error	0.14453					
df     SS     MS     F     Significance F       Regression     8     8.56023     1.07003     51.22610     2.13679E-53       Residual     329     6.87227     0.02089     2.13679E-53       Total     337     15.43250     0.02089     0.00000     8.69599       Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.00001     -0.27227     -0.10463	Observations	338					
Regression     8     8.56023     1.07003     51.22610     2.13679E-53       Residual     329     6.87227     0.02089     0.02089     0.02089       Total     337     15.43250     P-value     Lower 95%     Upper 95%       Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.000045     -0.11447     -0.03276       X1D3     -0.18845     0.04261     -4.42279     0.00001     -0.27227     -0.10463 </td <td>ANOVA</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td></td>	ANOVA						
Residual Total     329 337     6.87227 15.43250     0.02089       Coefficients     Standard Error     t Stat     P-value     Lower 95%     Upper 95%       Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.00045     -0.11447     -0.03276       X1D3     -0.18845     0.04261     -4.42279     0.00001     -0.27227     -0.10463		df	SS	MS	F	Significance F	
Total     337     15.43250       Coefficients     Standard Error     t Stat     P-value     Lower 95%     Upper 95%       Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.00045     -0.11447     -0.03276       X1D3     -0.18845     0.04261     -4.42279     0.00001     -0.27227     -0.10463	Regression	8	8.56023	1.07003	51.22610	2.13679E-53	
CoefficientsStandard Errort StatP-valueLower 95%Upper 95%Intercept8.810300.05811151.614650.000008.695998.92462X1-0.084370.01330-6.345910.00000-0.11052-0.05822X2-0.041270.00911-4.532100.00001-0.05919-0.02336D20.370000.111903.306610.001050.149870.59012D30.637810.170683.737010.000220.302060.97357D40.201470.132791.517170.13018-0.059760.46270X1D2-0.073620.02077-3.544490.00045-0.11447-0.03276X1D3-0.188450.04261-4.422790.00001-0.27227-0.10463	Residual	329	6.87227	0.02089			
Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.00045     -0.11447     -0.03276       X1D3     -0.18845     0.04261     -4.42279     0.00001     -0.27227     -0.10463	Total	337	15.43250				
Intercept     8.81030     0.05811     151.61465     0.00000     8.69599     8.92462       X1     -0.08437     0.01330     -6.34591     0.00000     -0.11052     -0.05822       X2     -0.04127     0.00911     -4.53210     0.00001     -0.05919     -0.02336       D2     0.37000     0.11190     3.30661     0.00105     0.14987     0.59012       D3     0.63781     0.17068     3.73701     0.00022     0.30206     0.97357       D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.00045     -0.11447     -0.03276       X1D3     -0.18845     0.04261     -4.42279     0.00001     -0.27227     -0.10463		0	<u>.</u>				
X1-0.084370.01330-6.345910.00000-0.11052-0.05822X2-0.041270.00911-4.532100.00001-0.05919-0.02336D20.370000.111903.306610.001050.149870.59012D30.637810.170683.737010.000220.302060.97357D40.201470.132791.517170.13018-0.059760.46270X1D2-0.073620.02077-3.544490.00045-0.11447-0.03276X1D3-0.188450.04261-4.422790.00001-0.27227-0.10463							
X2-0.041270.00911-4.532100.00001-0.05919-0.02336D20.370000.111903.306610.001050.149870.59012D30.637810.170683.737010.000220.302060.97357D40.201470.132791.517170.13018-0.059760.46270X1D2-0.073620.02077-3.544490.00045-0.11447-0.03276X1D3-0.188450.04261-4.422790.00001-0.27227-0.10463							
D20.370000.111903.306610.001050.149870.59012D30.637810.170683.737010.000220.302060.97357D40.201470.132791.517170.13018-0.059760.46270X1D2-0.073620.02077-3.544490.00045-0.11447-0.03276X1D3-0.188450.04261-4.422790.00001-0.27227-0.10463	X1	-0.08437	0.01330	-6.34591	0.00000	-0.11052	-0.05822
D30.637810.170683.737010.000220.302060.97357D40.201470.132791.517170.13018-0.059760.46270X1D2-0.073620.02077-3.544490.00045-0.11447-0.03276X1D3-0.188450.04261-4.422790.00001-0.27227-0.10463	X2	-0.04127	0.00911	-4.53210	0.00001	-0.05919	-0.02336
D4     0.20147     0.13279     1.51717     0.13018     -0.05976     0.46270       X1D2     -0.07362     0.02077     -3.54449     0.00045     -0.11447     -0.03276       X1D3     -0.18845     0.04261     -4.42279     0.00001     -0.27227     -0.10463	D2	0.37000	0.11190	3.30661	0.00105	0.14987	0.59012
X1D2     -0.07362     0.02077     -3.54449     0.00045     -0.11447     -0.03276       X1D3     -0.18845     0.04261     -4.42279     0.00001     -0.27227     -0.10463	D3	0.63781	0.17068	3.73701	0.00022	0.30206	0.97357
X1D3 -0.18845 0.04261 -4.42279 0.00001 -0.27227 -0.10463	D4	0.20147	0.13279	1.51717	0.13018	-0.05976	0.46270
	X1D2	-0.07362	0.02077	-3.54449	0.00045	-0.11447	-0.03276
X1D4 -0.07577 0.03254 -2.32835 0.02050 -0.13079 -0.01175	X1D3	-0.18845	0.04261	-4.42279	0.00001	-0.27227	-0.10463
	X1D4	-0.07577	0.03254	-2.32835	0.02050	-0.13979	-0.01175

Note that while  $R^2$  has slightly decreased compared with previous model, the adjusted  $R^2$  is higher than previous (54.39% vs 54.12%) and the F value has improved from 41 to 51. The only remaining questionable significance is that of  $D_4$  and to a lesser extent of  $X_1D_4$ . The poor p-values for these variables can be explained by a small amount of data in the West region (32 data points). Moreover, the

standard error of all dummy variables is relatively high, since these variables represent the addition to the intercept value, and there is a leverage effect of the slope error on the intercept error.

### Conclusion

The final model selected is determined by the following equation with coefficients from Table 9:

 $\mathbf{Y} = 8.8103 - 0.08437 \mathbf{X}_1 - 0.04127 \mathbf{X}_2 + 0.37 \mathbf{D}_2 + 0.63781 \mathbf{D}_3 + 0.20147 \mathbf{D}_4 - 0.07362 \mathbf{X}_1 \mathbf{D}_2 - 0.18845 \mathbf{X}_1 \mathbf{D}_3 - 0.07577 \mathbf{X}_1 \mathbf{D}_4,$ 

where Y is the natural logarithm of the price in Euros per square meter, and  $X_1$  is the square root of the distance from Luxembourg City in km.

Substituting the corresponding values of dummy variables for each region, the following equations are obtained:

Center / East:	$\mathbf{Y} = 8.81030 - 0.08437 \mathbf{X}_{1} - 0.04127 \mathbf{X}_{2}$
North :	$\mathbf{Y} = 9.18030 - 0.15799 \mathbf{X}_1 - 0.04127 \mathbf{X}_2$
South :	$\mathbf{Y} = 9.44811 - 0.27282\mathbf{X}_1 - 0.04127\mathbf{X}_2$
West :	$\mathbf{Y} = 9.01177 - 0.16014\mathbf{X}_1 - 0.04127\mathbf{X}_2$

Note that intercept and slope coefficients for  $X_1$  compare well with those of the single variable regressions in Chart 4.

# Appendix – Original data, excl. Luxembourg City, sorted by distance

Thhe			igi					embo	urg			rted by		Stanc	6		
Price, EUR	Area, m2	Bedrooms	Туре	Distance km	Region	Price, EUR	Area, m2	Bedrooms	Туре	Distance km	, Region	Price, EUR	Area, m2	Bedrooms	Туре	Distance km	Region
524,713 592,430	128 157	2 3	A A	4 4	Center Center	610,470 251,339	139 49	3 1	H A	16 16	South South	536,952 449.813	130 146	4 4	H H	26 26	North North
588,608	168	4	А	4	Center	387,569	81	2	А	16	South	433,803	94	2	А	26	North
495,000 612,900	105 92	2 1	A A	4	Center Center	488,680 476,890	104 114	3 3	A A	16 16	South South	388,500 571,842	87 113	2 3	A A	26 26	North North
457,950	82	2	А	4	Center	248,146	54	1	А	16	South	408,952	87	2	А	26	North
575,900 971,006	104 173	3 4	A H	4 5	Center Center	390,303 558,400	85 116	2 3	A A	16 16	South South	618,281 466,000	125 139	3 3	A H	26 26	North North
1,017,500	180	4	н	5	Center	187,831	45	1	A	16	South	286,501	61	1	A	26	North
272,500 377,500	53 75	1 2	A A	6 6	Center Center	316,450 385,511	83 96	2 3	A A	16 16	South South	738,295 584,096	144 133	2 3	н Н	27 27	East East
487,500 551,240	94 121	3	A	6 6	Center Center	598,905 284,604	143 59	4	A A	16 16	South South	490,119 309,365	90 76	1 2	A A	27 27	East East
338,236	63	1	Â	6	Center	434,293	79	2	A	16	South	385,450	107	3	A	27	East
611,961 515,873	128 123	3 2	A A	6 6	Center Center	214,900 349,250	46 79	1 2	A A	16 16	South South	312,922 191,400	106 55	2	A A	27 27	East East
520,966	124	3	А	6	Center	444,100	105	3	A	16	South	149,800	44	1	A	27	East
393,592 481,454	89 111	2 3	A A	6 6	Center Center	364,068 349,808	101 97	2 3	A A	16 16	South South	160,389 401,175	44 112	2 3	A A	27 27	East East
312,642	61	1	А	6	Center	459,146	134	3	A	16	South	235,939	59	1	A	27	West
565,642 853,450	117 173	3	A H	6 7	Center Center	377,282 365,608	127 85	2 2	A A	16 16	South South	424,230 425,000	126 129	3 3	A A	27 28	West North
461,900	90	2	A A	7 7	South	745,000	234 170	4 3	H H	16	East	431,430	165 174	4	н н	28	North
539,500 730,100	114 137	3 4	A	7	South South	598,988 673,042	186	3	Н	16 16	East East	673,116 330,637	90	3 2	A	28 28	North North
476,000 791,343	93 225	2 3	A H	8 8	West West	530,000 471,203	140 160	3 3	H H	17 17	East West	406,509 590,000	118 130	3 3	A H	28 28	North North
791,343	174	3	Н	8	West	380,100	88	2	A	17	West	495,000	162	4	н	29	North
529,400 556,000	105 124	2 3	H A	8 8	West West	485,400 846,042	114 221	3 3	A H	17 17	West West	516,600 404,754	123 100	3 1	A A	29 29	North North
437,694	83	2	А	9	Center	211,520	53	1	A	17	West	434,861	107	2	А	29	North
476,650 968,347	96 214	3	A H	9 9	Center Center	263,600 461,805	68 99	2 2	A A	17 18	West East	427,308 539,320	115 143	3 3	A A	29 29	North North
777,300	140	3	н	9	Center	411,427	85	2	Α	18	East	558,000	134	3	н	29	East
1,268,673 979,076	198 198	3 4	н Н	10 10	Center Center	1,300,000 532,265	367 146	3 1	A A	18 18	East East	596,150 603,881	141 176	4 3	H H	29 29	East East
709,420	150	3	н	10	Center	461,805	98	2	A	18	East	320,000	91	2	А	30	North
562,000 272,899	114 63	3 2	H A	10 10	Center Center	512,550 279,913	99 76	3 2	A A	18 18	East West	827,760 255,971	142 58	4 1	H A	30 30	East East
427,721	99	3	А	10	Center	316,838	83	2	A	18	West	336,253	106	2	A	32	North
887,581 423,183	180 92	4	H A	10 10	Center South	376,800 375,000	98 98	3 2	A A	18 19	West Center	414,087 631,083	141 165	3 3	H H	32 32	West West
609,196	131 157	3 4	н н	10 10	South West	268,000 445,270	84 106	2 3	A A	19 19	South South	554,940 305,750	150 75	3 2	H A	33 33	North North
840,000 268,021	51	4	A	10	West	306,000	95	2	A	19	South	395,000	144	2	A	34	East
628,025 1,052,940	162 198	3 3	H H	10 10	West West	200,000 215,000	66 65	1 2	A A	19 19	South South	672,500 229,188	175 58	4 1	H A	34 36	East North
440,166	101	2	А	11	South	329,000	125	3	А	19	South	319,122	91	2	А	36	North
611,950 590,700	151 149	3 3	A A	11 11	South South	231,156 256,796	54 56	1 2	A A	19 19	South South	218,000 220,000	77 77	1 2	A A	36 36	North North
655,495	160	4	н	11	West	175,000	56	1	A	19	South	238,000	86	1	А	36	North
750,000 469,000	184 108	4	H A	11 11	West West	230,000 245,000	73 77	2 2	A A	19 19	South South	302,000 327,000	101 113	2 3	A A	36 36	North North
334,806	65	1	А	12	Center	320,200	100	3	A	19	South	416,000	120	3	н	36	North
407,892 492,352	77 102	2	A	12 12	Center Center	164,164 215,000	53 71	1 2	A A	19 19	South South	351,147 598,740	145 180	4 3	H H	36 36	North North
439,500	91	2	А	12	Center	243,984	58	1	A	19	South	206,353	71	2	А	36	North
1,035,325 310,000	231 73	3 2	H A	12 14	Center South	291,225 244,800	86 61	2 1	A A	19 19	South South	250,806 460,000	79 147	3 3	A H	36 36	North North
413,080	83	2	A	14	South	198,090	63	1	A	19	South	272,000	71	1	A	36	North
474,369 296,160	104 70	3 1	A A	14 14	South South	257,108 259,000	83 51	2 1	A A	19 19	South East	262,000 391,000	78 124	2 3	A A	36 36	North North
356,952	73	2	A	14	South	350,400 452,300	78 109	2 3	A A	19 19	East East	453,000 407,344	130 163	4 3	H H	38 38	North North
419,346 227,622	104 56	3 1	A A	14 14	South South	787,139	176	3	н	20	Center	463,283	135	4	н	38	North
328,769 256,392	78 71	2 2	A A	14 14	South South	816,956 276,000	220 82	4	H A	20 20	Center South	273,201 287,636	76 85	1 2	A A	42 42	North North
372,000	85	2	Â	15	Center	162,180	51	1	А	20	South	457,079	126	3	А	42	North
732,402 488,000	145 114	3	н Н	15 15	Center Center	333,020 380,000	103 126	3	A A	20 20	South South	270,799 198,609	90 62	2 1	A A	42 42	North North
593,000	130	3	н	15	Center	289,000	82	2	A	20	South	284,147	93	2	А	42	North
743,910 539,000	173 140	3 4	н Н	15 15	Center Center	435,000 364,902	144 105	3 3	H A	20 20	South South	235,000 245,000	82 82	1 2	A A	42 42	North North
428,227	97	2	А	15	Center	682,645	151	3	н	20	East	232,154	89	2	А	42	North
465,199 460,587	109 106	2 2	A A	15 15	Center Center	550,960 334,890	147 107	3 3	H A	21 21	Center East	180,713 268,268	70 103	2 3	A A	42 42	North North
395,850	93	3	А	15	Center	477,324	190	3	н	21	West	248,000	67	2	A	43	North
501,000 448,000	150 110	3 3	н Н	15 15	Center Center	295,336 368,282	83 92	1 2	A A	23 23	Center Center	316,000 440,000	101 135	3 4	A H	43 43	North North
595,000 598,000	114 140	3 3	н Н	15 15	Center Center	381,156 495,000	109 127	3 3	A A	23 23	Center Center	204,618 343,300	87 136	2 3	A A	44 44	North North
394,000	100	3	н	15	Center	272,991	67	1	Α	23	North	310,000	129	3	н	44	North
458,587 732,402	103 155	2	A H	15 15	Center Center	326,745 658,005	85 187	2 4	A H	23 24	North East	482,476 497,248	140 170	3 3	H H	45 45	North North
698,000	175	4	н	15	Center	349,000	79	2	Α	24	East	442,510	174	4	н	45	North
539,000 405,000	140 91	4	н Н	15 15	Center South	418,000 660,000	127 130	3 3	A H	24 24	East East	680,500 216,000	220 80	4	H A	49 50	North North
310,000	80	2	А	15	South	768,500	175	4	н	24	East	246,000	94	3	А	50	North
638,000 311,261	125 49	3 1	H A	15 15	East East	410,000 610,440	97 227	2 4	A H	24 24	East East	272,079 272,710	75 89	1 2	A A	50 50	North North
431,562	77	2	A	15	East	485,000	160	3	н	24	East	346,736	116	3	A	50	North
549,293 227,130	103 54	3 1	A H	15 15	East East	316,800 411,600	86 115	2 3	A A	24 24	East East	446,050 298,400	166 84	3 2	H A	50 50	North North
375,970	94 108	2 3	н н	15 15	East East	478,000	120 152	3 3	H H	24 24	West West	445,300	142 84	4	A A	50 50	North North
425,240 469,738	111	3	A	15	East	565,742 628,000	120	3	Н	24	North	268,100 408,100	04 140	4	A	50	North
350,000	101 180	3	A H	15 15	East	667,000 363,200	155 84	4	H A	25 25	North North	197,060 275,000	67 90	2	A H	51 51	North North
706,017 570,735	180 157	3	н	15 15	West West	339,600	77	2	А	25	North	354,716	130	3	н	51	North
470,100 743,094	102	3 4	н н	16 16	Center	459,000 480,775	111 230	3 4	A H	25 25	North West	560,143 342,504	160 123	3 4	H A	51 51	North North
685,000	171 192	4	н Н	16 16	Center South	402,420	160	3	н	25	West	129,049	57	1	А	57	North
498,000 237,736	137 57	3 1	A A	16 16	South South	359,500 436,100	77 97	1 2	A A	26 26	North North	208,816 214,699	97 99	2 3	A A	57 57	North North
393,460	93	2	А	16	South	228,000	54	2	A	26	North	434,344	176	4	н	57	North
293,790 303,240	67 73	1	A A	16 16	South South	685,557 296,900	187 79	4	A A	26 26	North North	253,104 274,623	77 82	1 2	A A	57 57	North North
349,808	114	3	А	16	South	380,668	108	3	А	26	North	111,240	37	1	A	59	North
320,290 319,076	91 79	1 2	A A	16 16	South South	245,000 498,600	69 130	2 3	A H	26 26	North North	221,370 326,490	74 109	2 3	A A	59 59	North North
488,593	150	3	А	16	South	472,242	118	3	н	26	North	180,259	59	1	А	59	North
285,840 493,530	77 124	2 3	A	16 16	South South	513,820 524,191	143 136	3 3	H H	26 26	North North	232,880	79	2	A	59	North