

## ÚKOL 9

### CVIČENÍ Z ANALÝZY DAT SOCIOLOGICKÉHO VÝZKUMU

Jméno a příjmení:

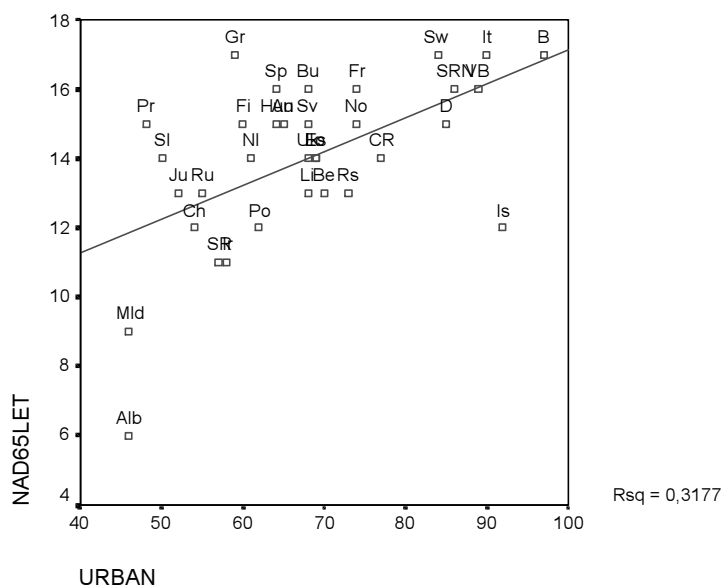
Imatrik. ročník:

1.

V datech dmg-data vypočítejte regresní rovnici pro vztah mezi podílem osob ve věku nad 65 let a mírou urbanizace (předpokládáme, že s mírou urbanizace země, která je indikátorem modernity, se bude zvyšovat podíl osob starších 65 let). Zjistěte, zdali má smysl pracovat s celým souborem, nebo zdá má cenu uvažovat pouze o nějakém podsouboru zemí.

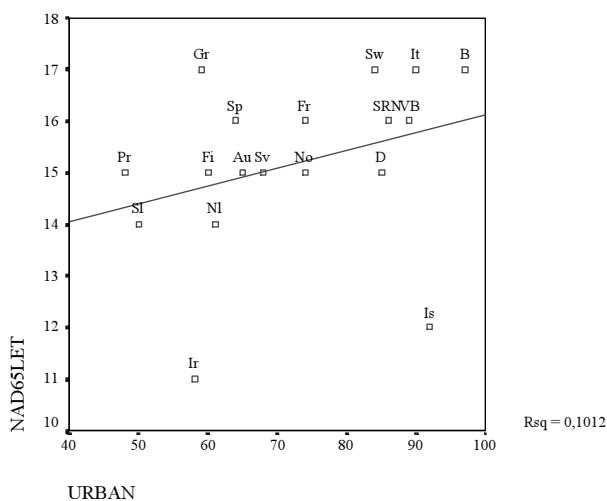
Řešení:

Nejdříve zjistíme graficky, zdali je vztah lineární povahy:  
Scatter graph



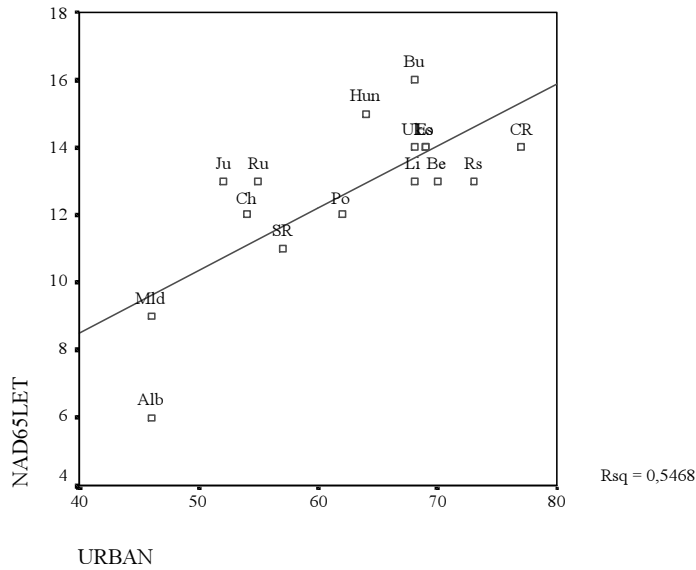
Linearita je v souboru všech zemí slabá (Rsq je pouze 0,32).

**Podsoubor západních zemí:**



Linearita velmi nízká.

**Podsoubor východních zemí:**



Linearita slušná.

**Regrese:**

Analyze-Regression-Linear

**Model Summary<sup>b</sup>**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,739 <sup>a</sup>	,547	,514	1,6660

a. Predictors: (Constant), URBAN

b. Dependent Variable: NAD65LET

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	46,891	1	46,891	16,894	,001 <sup>a</sup>
	Residual	38,859	14	2,776		
	Total	85,750	15			

a. Predictors: (Constant), URBAN

b. Dependent Variable: NAD65LET

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1,076	2,841		,379	,711
	URBAN	,185	,045	,739	4,110	,001

a. Dependent Variable: NAD65LET

Rovnice:

$$\text{podíl nad 65 let} = 1,08 + 0,185 \cdot \text{urbanizace}$$

Zvýší-li se podíl lidí žijících ve městech o 1 procento, zvýší se podíl osob starších 65 let o 0,19%.

## 2.

Vyjděte z příkladu 1 a zjistěte, které z východoevropských zemí se ve svých predikovaných hodnotách nejvíce liší od hodnot skutečných?

Case Summaries <sup>a</sup>

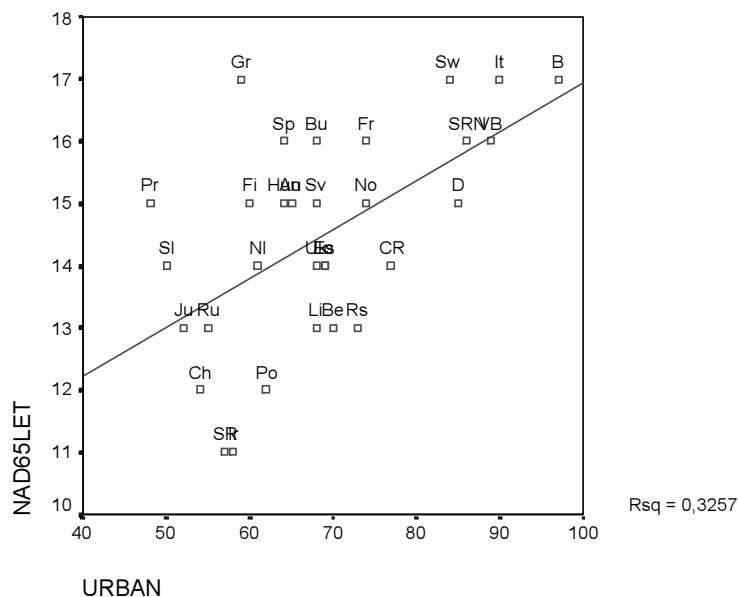
	ZEME Udaje za rok 1999 (population data sheet)	NAD65LET	PRE_2 Unstandardized Predicted Value	RES_2 Unstandardized Residual
1	Albanie	6,00	9,59304	-3,59304
2	Belorusko	13,00	14,03683	-1,03683
3	Bulharsko	16,00	13,66651	2,33349
4	Cesko	14,00	15,33294	-1,33294
5	Estonsko	14,00	13,85167	,14833
6	Chorvatsko	12,00	11,07430	,92570
7	Jugoslavie	13,00	10,70398	2,29602
8	Litevsko	13,00	13,66651	-,66651
9	Lotyšsko	14,00	13,85167	,14833
10	Maďarsko	15,00	12,92588	2,07412
11	Moldavie	9,00	9,59304	-,59304
12	Polsko	12,00	12,55557	-,55557
13	Rumunsko	13,00	11,25946	1,74054
14	Rusko	13,00	14,59230	-1,59230
15	Slovensko	11,00	11,62978	-,62978
16	Ukrajina	14,00	13,66651	,33349
Total	N	16	16	16

<sup>a</sup>. Limited to first 100 cases.

Nejvíce se liší Albánie, kde je podíl lidí nad 65 o 3,5 % nižší než predikovaný podle našeho regresního modelu. Dále se pak hodně liší Bulharsko, Jugoslávie a Maďarsko.

## 3.

Jak se změní obraz linearity, když z celého souboru evropských zemí vyloučíte největší outliers: Albánii, Moldávii a Island?



Ve srovnání se souborem všech zemí téměř nijak. R<sub>sq</sub> je pořád nízké, 0,33 (pro celý soubor bylo 0,32).