

# Metodologický seminář – III. JASP - Lineární regresní analýza

Pavlína Janošová

[janosova@htf.cuni.cz](mailto:janosova@htf.cuni.cz)

Konzultace **po dohodě**:

v MS Teams

případně přes Skype

či Zoom



## Informace k lineární regresi v Excelu (Youtube)

**Petr Soukup – seminář o jednoduché lineární regresi v Excelu (česky):**

<https://www.youtube.com/watch?v=r28x1NYeiyA>

Informace o výpočtu jednoduché lineární regrese v Excelu. **DŮLEŽITÁ ČÁST JE OD 4. MINUTY PŘIBLIŽNĚ DO KONCE** (konci rozumět nemusíte). **Cca 52 minut, tj. za 55 minut sraz v MS Teamsech.**

**Petr Soukup – seminář o jednoduché lineární regresi v JASP (anglicky):**

<https://www.youtube.com/watch?v=Sxuc1Vvxidk>

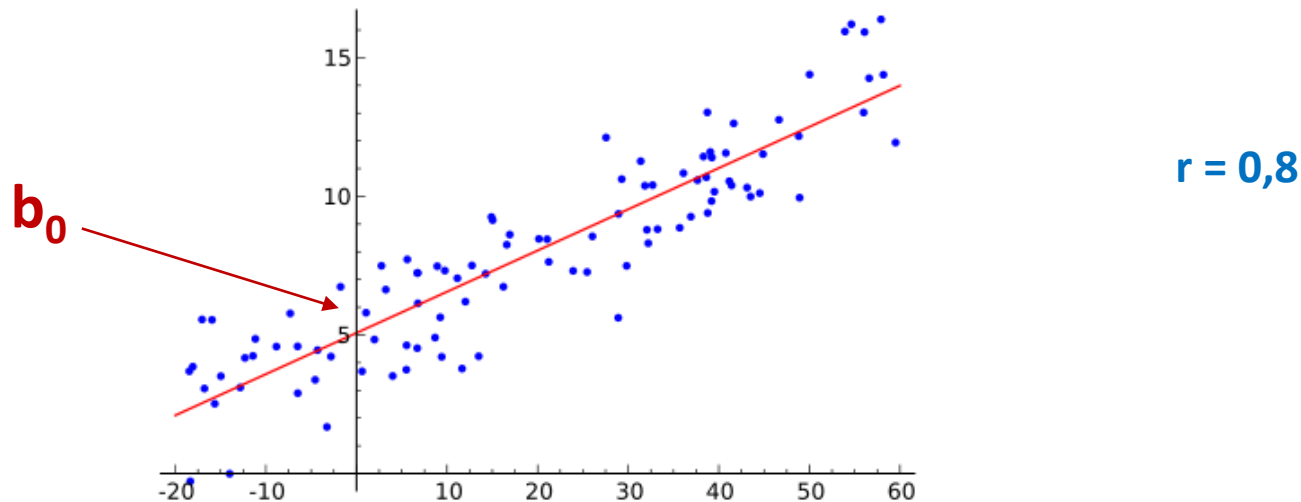
Tentýž typ výpočtu jako v předchozím videu, avšak v JASP. Základní informace si řekneme v prezentaci.

*Petr Soukup – seminář o vícenásobné lineární regresi ve statistickém programu Mplus (anglicky)*

<https://www.youtube.com/watch?v=xLlvoCY6oC4&t=2560s>

## Informace k lineární regresi v Excelu (Youtube)

U lineární regrese předpokládáme lineární vztah mezi proměnnými, vyjádřitelný **přímkou**.



$$Y = b_0 + b_1 \cdot x$$

$b_0$  je konstanta, tj. průsečík přímky s osou  $y$  (zde  $b_0 = 5$ )

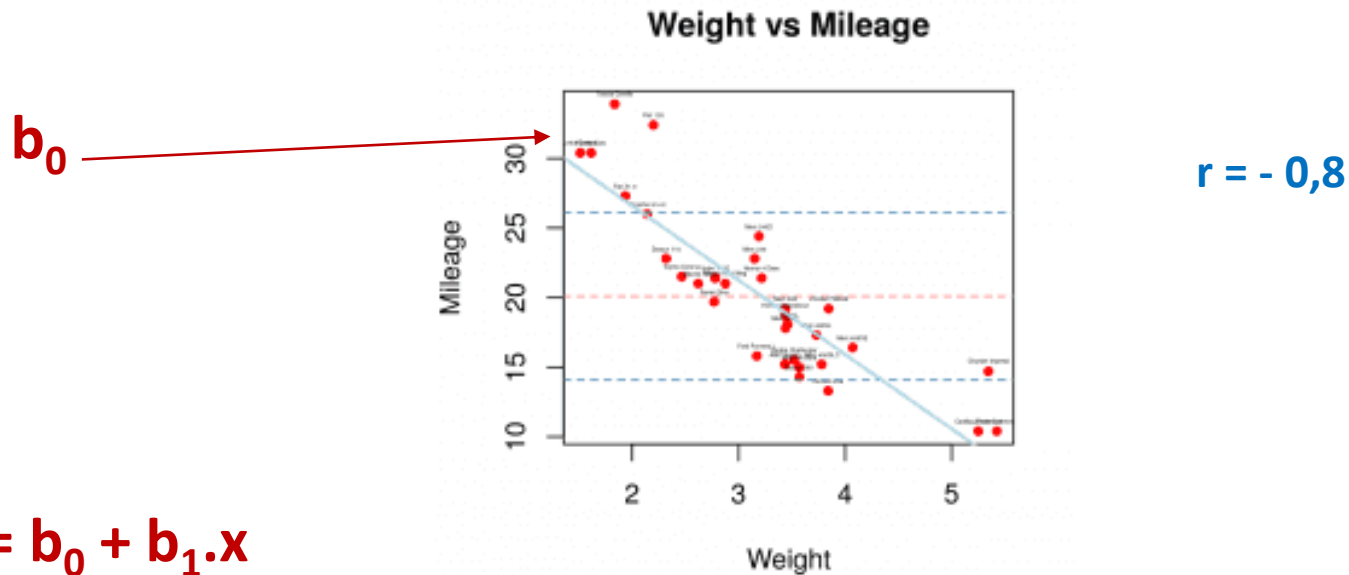
$b_1$  je informace o směru a sklonu přímky (= nahoru... dolů... rovně)

⇒ s růstem hodnoty „ $x$ “ nějak rostou (nebo klesají) i hodnoty na ose „ $y$ “

⇒ kladná hodnota  $b_1$  = přímá úměrnost

## Informace k lineární regresi v Excelu (Youtube)

U lineární regrese předpokládáme lineární vztah mezi proměnnými, vyjádřitelný **přímkou**.



$$Y = b_0 + b_1 \cdot x$$

$b_0$  je konstanta, tj. průsečík přímky s osou  $y$

$b_1$  je informace o směru a sklonu přímky (= nahoru... dolů... rovně)

⇒ s růstem hodnoty „ $x$ “ klesá i hodnota na ose „ $y$ “

⇒ záporná hodnota  $b_1$  = nepřímá úměrnost



## Základní informace o vícenásobné regresní analýze

Pomocí lineární regresní analýzy můžeme odhadovat hodnotu závisle proměnné i z více než jedné nezávisle proměnné. Takové modely se více podobají reálnému životu, protože téměř nikdy na nějaký jev nepůsobí pouze jeden vliv.

Zjednodušeně si to můžeme představit i jako vícenásobnou korelaci závisle proměnné s proměnnými nezávislými.

*VZOREC VÍCENÁSOBNÉ REGRESNÍ ANALÝZY:*

$$Y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n$$

Počátek/konstanta/intercept

jednotlivé nezávislé proměnné

$b_1, b_2, \dots, b_n$  – velikosti sklonů u těchto proměnných

## Krátká ukázka vícenásobné regresní analýzy



Výzkumná otázka – **jaký je odhad, že žák získá empatii od svých spolužáků v závislosti na tom, (1) zda je šikanovaný?**

Máme tři kontrolní proměnné: (2) jak je sám empatický; (3) postavení tohoto žáka ve třídě a (4) pohlaví.

**Pohlaví**

**Empatie k druhým (sebeuposouzení)**

**Skupinový status (vrstevnické nominace)**

**Viktimizace (sebeuposouzení)**

**získaná empatie**

**od spolužáků**

*Závisle proměnná by měla vzniknout v časové souslednosti až po vzniku proměnných nezávislých*

# Krátká ukázka vícenásobné regresní analýzy



## (1) Kontrola popisných dat u hlavních proměnných:

**Olw2** => Olweusův sebeposuzovací dotazník (jak často byli v posledních měsících šikanováni; 1 až 5, tj. pořadová/ordinální proměnná. Její úpravou na spojitou/kardinální proměnnou dost „švindlujeme“).

**GW\_KOGNIT. EMPAT** => získaná empatie od spolužáků (kdo se ve třídě necítí dobře, ke komu jsou ostatní ve třídě nespravedliví).

The screenshot shows the SPSS software interface. The top menu bar includes options like Descriptives, T-Tests, ANOVA, Mixed Models, Regression, Frequencies, Factor, Meta-Analysis, and Reliability. The main window is divided into three panes:

- Left Pane (Variable List):** Lists variables such as empat.effic1-6, Soc.effic1-5, empatie1-4, Olw1, and Olw3. Olw3 is selected.
- Middle Pane (Variables):** Shows the selected variables GW\_KOGNIT.EMPAT and Olw2.
- Right Pane (Results):** Displays the Descriptive Statistics for the selected variables.

	GW_KOGNIT.EMPAT	Olw2
Valid	184	184
Missing	0	0
Mean	0.049	1.473
Median	0.000	1.000
Std. Deviation	0.112	0.923
Minimum	0.000	1.000
Maximum	0.632	5.000

# Krátká ukázka vícenásobné regresní analýzy



## (1) Kontrola popisných dat u hlavních proměnných

**Olw2** => Olweusův sebeposuzovací dotazník (jak často byli v posledních měsících šikanováni; 1 až 5)

**GW\_KOGNIT. EMPAT** => získaná empatie od spolužáků (kdo se ve třídě necítí dobře, nespravedliví)

The screenshot shows the SPSS software interface. The top menu bar includes Descriptives, T-Tests, ANOVA, Mixed Models, Regression, Frequencies, Factor, Meta-Analysis, and Reliability. The Descriptives dialog box is open, showing the following options checked:

- Central Tendency: Mean, Median
- Dispersion: Standard deviation, Minimum, Maximum

The Results window displays the following table:

	GW_KOGNIT.EMPAT	Olw2
Valid	184	184
Missing	0	0
Mean	0.049	1.473
Median	0.000	1.000
Std. Deviation	0.112	0.923
Minimum	0.000	1.000
Maximum	0.632	5.000

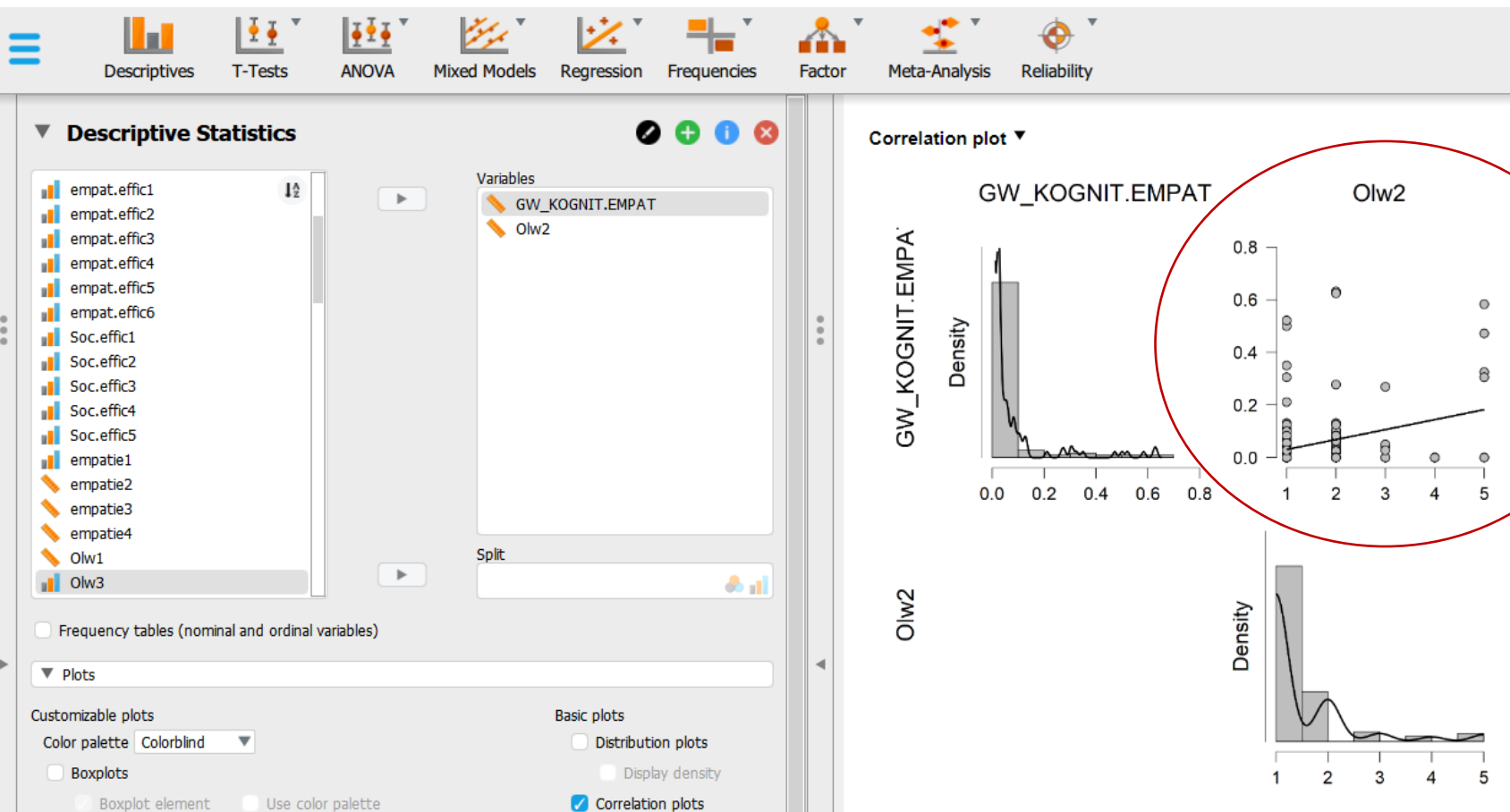


# Krátká ukázka vícenásobné regresní analýzy



**(1) Kontrola linearity vztahu mezi hlavními proměnnými (zde je slabá natolik, že si tento výpočet nemůžeme v reálu dovolit, ale abychom si vyzkoušeli tento typ výpočtu, tak ji provedeme). Kontrolujeme i odlehlé hodnoty.**

*- S narůstající četností viktimizace mírně narůstá i empatie spolužáků vůči němu.*



# Krátká ukázka vícenásobné regresní analýzy



(1) Kontrola **kolinearity** vztahu mezi nezávisle proměnnými (je nízká, takže tuto podmínku splňujeme).

The screenshot displays the SPSS software interface. The top menu bar includes Descriptives, T-Tests, ANOVA, Mixed Models, Regression, Frequencies, Factor, Meta-Analysis, and Reliability. The main window is divided into two panes. The left pane, titled "Linear Regression", contains a "Correlation" dialog box. The "Variables" list includes Olw2, self\_AFEKT.EMPAT, and MEAN\_STATUS\_NEW. The "Condition on" list is empty. The "Sample Correlation Coefficient" section has "Pearson's r" checked. The "Additional Options" section has "Display pairwise" and "Report significance" checked. The right pane, titled "Results", shows a "Correlation" section with a "Pearson's Correlations" table. The table has three columns: the first column lists the variables, the second column lists the variables they are correlated with, the third column is "Pearson's r", and the fourth column is "p". The value 0.110 in the "Pearson's r" column for the correlation between Olw2 and self\_AFEKT.EMPAT is circled in red, with a red arrow pointing to it from the right side of the slide.

		Pearson's r	p
Olw2	- self_AFEKT.EMPAT	0.110	0.137
Olw2	- MEAN_STATUS_NEW	-0.225	0.002
self_AFEKT.EMPAT	- MEAN_STATUS_NEW	0.032	0.670

# VÝPOČET vícenásobné regresní analýzy - I



## Regression => Classical => Linear Regression

*Sledujeme, zda je model vůbec statisticky významný, zda se alespoň některé nezávisle proměnné významně podílejí na vysvětlení závisle proměnné.*

The screenshot displays the SPSS Linear Regression dialog box and the Results window. In the dialog box, the dependent variable is 'GW\_KOGNIT.EMPAT' and the method is 'Enter'. Covariates include 'self\_AFEKT.EMPAT', 'pohlavi', 'MEAN\_STATUS\_NEW', and 'Olw2'. The 'Model fit' checkbox is checked. In the Results window, the ANOVA table shows a significant F-value of 10.260 (p < .001).

**Linear Regression**

Dependent Variable: GW\_KOGNIT.EMPAT

Method: Enter

Covariates: self\_AFEKT.EMPAT, pohlavi, MEAN\_STATUS\_NEW, Olw2

WLS Weights (optional):

Model fit:

Statistics:  Model fit,  Descriptives

**Results**

**Linear Regression**

Model Summary - GW\_KOGNIT.EMPAT

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	RMSE
H <sub>0</sub>	0.000	0.000	0.000	0.112
H <sub>1</sub>	0.432	0.187	0.168	0.102

ANOVA

Model		Sum of Squares	df	Mean Square	F	p
H <sub>1</sub>	Regression	0.428	4	0.107	10.260	< .001
	Residual	1.865	179	0.010		
	Total	2.293	183			

Note. The intercept model is omitted, as no meaningful information can be shown.

# VÝPOČET vícenásobné regresní analýzy - II



Sledujeme, **kolik procent rozptylu závisle proměnné** tyto nezávisle proměnné **vysvětlují**.

- 17 % vysvětleného rozptylu znamená, že závisle proměnná souvisí mnohem víc s jinými vlivy než s těmi, které zkoumáme.

The screenshot shows the SPSS Linear Regression dialog box on the left and the Results window on the right. The dialog box has 'GW\_KOGNIT.EMPAT' as the dependent variable and 'self\_AFEKT.EMPAT', 'pohlavi', 'MEAN\_STATUS\_NEW', and 'Olw2' as covariates. The 'Enter' method is selected. The Results window displays the Model Summary and Coefficients tables.

### Linear Regression

Dependent Variable: GW\_KOGNIT.EMPAT  
Method: Enter

Covariates: self\_AFEKT.EMPAT, pohlavi, MEAN\_STATUS\_NEW, Olw2

### Results

#### Linear Regression

Model Summary - GW\_KOGNIT.EMPAT

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	FMSE
H <sub>0</sub>	0.000	0.000	0.000	0.112
H <sub>1</sub>	0.432	0.187	0.168	0.102

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
H <sub>0</sub>	(Intercept)	0.049	0.008		5.881	< .001
H <sub>1</sub>	(Intercept)	0.039	0.044		0.874	0.383
	self_AFEKT.EMPAT	-0.004	0.011	-0.028	-0.373	0.709
	pohlavi	0.030	0.016	0.134	1.818	0.071
	MEAN_STATUS_NEW	-0.222	0.060	-0.258	-3.725	< .001
	Olw2	0.032	0.008	0.266	3.812	< .001

#### Descriptives

	N	Mean	SD	SE
GW_KOGNIT.EMPAT	184	0.049	0.112	0.008
self_AFEKT.EMPAT	184	3.351	0.747	0.055
pohlavi	184	0.516	0.501	0.037
MEAN_STATUS_NEW	184	0.177	0.130	0.010
Olw2	184	1.473	0.923	0.068

# VÝPOČET vícenásobné regresní analýzy - III



**Intercept (počátek/konstanta)** je míra získané empatie od spolužáků, pokud jsou všechny hodnoty závisle proměnných rovny nule.

- Je to nereálná hodnota, protože míra viktimizace začínala např. na hodnotě 1 a sebehodnocení v empatii na hodnotě 4...

The screenshot displays the SPSS Linear Regression dialog box on the left and the Results window on the right.

**Linear Regression Dialog:**

- Dependent Variable: GW\_KOGNIT.EMPAT
- Method: Enter
- Covariates: self\_AFEKT.EMPAT, pohlavi, MEAN\_STATUS\_NEW, Olw2
- WLS Weights (optional):

**Results - Linear Regression**

Model Summary - GW\_KOGNIT.EMPAT

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	RMSE
H <sub>0</sub>	0.000	0.000	0.000	0.112
H <sub>1</sub>	0.432	0.187	0.168	0.102

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
H <sub>0</sub>	(Intercept)	0.049	0.008		5.881	< .001
H <sub>1</sub>	(Intercept)	0.039	0.044		0.874	0.383
	self_AFEKT.EMPAT	-0.004	0.011	-0.028	-0.373	0.709
	po hlavi	0.030	0.016	0.134	1.818	0.071
	MEAN_STATUS_NEW	-0.222	0.060	-0.258	-3.725	< .001
	Olw2	0.032	0.008	0.266	3.812	< .001

Descriptives

	N	Mean	SD	SE
GW_KOGNIT.EMPAT	184	0.049	0.112	0.008
self_AFEKT.EMPAT	184	3.351	0.747	0.055
po hlavi	184	0.516	0.501	0.037
MEAN_STATUS_NEW	184	0.177	0.130	0.010
Olw2	184	1.473	0.923	0.068



# VÝPOČET vícenásobné regresní analýzy - VI

Sledujeme, které nezávisle proměnné se významně podílejí na vysvětlení závisle proměnné:

The screenshot shows the SPSS Linear Regression dialog box on the left and the Results window on the right. The dialog box has 'GW\_KOGNIT.EMPAT' as the dependent variable and 'self\_AFEKT.EMPAT', 'pohlavi', 'MEAN\_STATUS\_NEW', and 'Olw2' as covariates. The Results window shows the Model Summary and Coefficients tables. A red circle highlights the p-values for the independent variables in the Coefficients table, with a red arrow pointing to the p-value for 'self\_AFEKT.EMPAT'.

**Linear Regression**

Dependent Variable: GW\_KOGNIT.EMPAT

Method: Enter

Covariates: self\_AFEKT.EMPAT, pohlavi, MEAN\_STATUS\_NEW, Olw2

WLS Weights (optional):

Model

Statistics

Regression Coefficients

- Estimates
- From 5000 bootstraps
- Confidence intervals 95.0 %
- Covariance matrix
- Vovk-Sellke maximum p-ratio
- Model fit
- R squared change
- Descriptives
- Part and partial correlations
- Collinearity diagnostics

**Results**

**Linear Regression**

Model Summary - GW\_KOGNIT.EMPAT

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	RMSE
H <sub>0</sub>	0.000	0.000	0.000	0.112
H <sub>1</sub>	0.432	0.187	0.168	0.102

Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
H <sub>0</sub>	(Intercept)	0.049	0.008		5.881	< .001
H <sub>1</sub>	(Intercept)	0.039	0.044		0.874	0.383
	self_AFEKT.EMPAT	-0.004	0.011	-0.028	-0.373	0.709
	pohlavi	0.030	0.016	0.134	1.818	0.071
	MEAN_STATUS_NEW	-0.222	0.060	-0.258	-3.725	< .001
	Olw2	0.032	0.008	0.266	3.812	< .001

Descriptives

	N	Mean	SD	SE
GW_KOGNIT.EMPAT	184	0.049	0.112	0.008
self_AFEKT.EMPAT	184	3.351	0.747	0.055
pohlavi	184	0.516	0.501	0.037
MEAN_STATUS_NEW	184	0.177	0.130	0.010
Olw2	184	1.473	0.923	0.068

# VÝPOČET vícenásobné regresní analýzy - VI



**Které nezávisle proměnné se významně podílejí na vysvětlení závisle proměnné?**

- *žáci, kteří jsou podle svých slov šikanováni, dostávají významně víc empatie (M. Kolář - ???)*
- *Čím nižší je skupinový status žáků, tím více empatie získávají (M. Kolář - ???)*
- *Získaná empatie od spolužáků nesouvisí s tím, zda jsou sami žáci vůči nim empatičtí (reciprocita ???)*

The screenshot displays the SPSS Linear Regression dialog box on the left and the Results window on the right. The dialog box shows the dependent variable 'GW\_KOGNIT.EMPAT' and covariates 'self\_AFEKT.EMPAT', 'pohlavi', 'MEAN\_STATUS\_NEW', and 'Olw2'. The Results window shows the Model Summary and Coefficients tables.

### Linear Regression

Dependent Variable: GW\_KOGNIT.EMPAT  
Method: Enter

Covariates: self\_AFEKT.EMPAT, pohlavi, MEAN\_STATUS\_NEW, Olw2

### Results

#### Linear Regression

Model Summary - GW\_KOGNIT.EMPAT

Model	R	R <sup>2</sup>	Adjusted R <sup>2</sup>	RMSE
H <sub>0</sub>	0.000	0.000	0.000	0.112
H <sub>1</sub>	0.432	0.187	0.168	0.102

#### Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p
H <sub>0</sub>	(Intercept)	0.049	0.008		5.881	< .001
H <sub>1</sub>	(Intercept)	0.039	0.044		0.874	0.383
	self_AFEKT.EMPAT	-0.004	0.011	-0.028	-0.373	0.709
	pohlavi	0.030	0.016	0.134	1.818	0.071
	MEAN_STATUS_NEW	-0.222	0.060	-0.258	-3.725	< .001
	Olw2	0.032	0.008	0.266	3.812	< .001

#### Descriptives

	N	Mean	SD	SE
GW_KOGNIT.EMPAT	184	0.049	0.112	0.008
self_AFEKT.EMPAT	184	3.351	0.747	0.055
pohlavi	184	0.516	0.501	0.037
MEAN_STATUS_NEW	184	0.177	0.130	0.010
Olw2	184	1.473	0.923	0.068



## VÝSLEDKY binární **logistické** analýzy

Předchozí model neodpovídal možnostem použití lineárního regresního modelu.

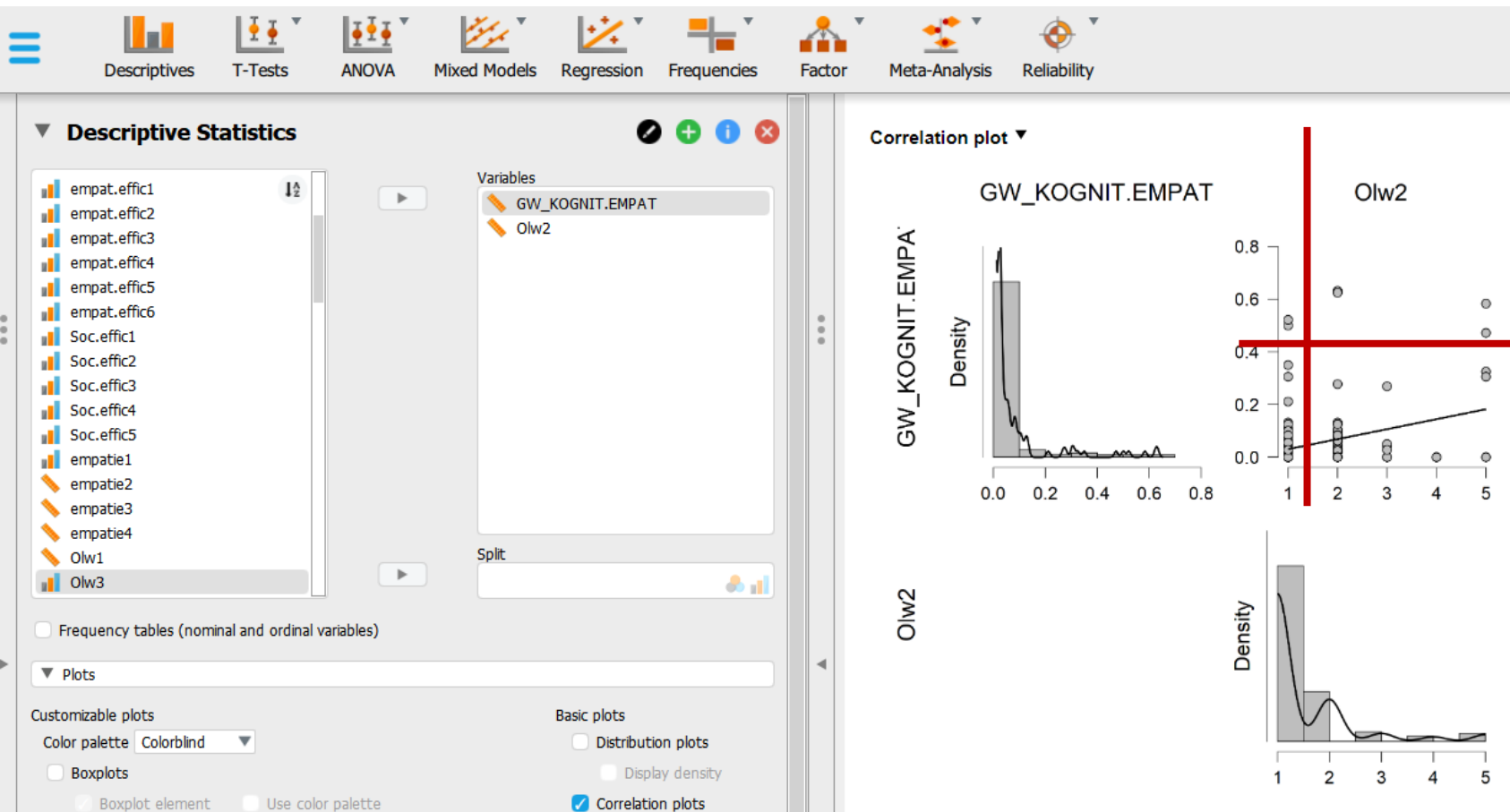
- Tuto výzkumnou otázku jsem řešila pomocí logistické regresní analýzy (= **kteří spolužáci získají nadprůměrnou míru empatie od spolužáků?**)



# VÝSLEDKY binární **logistické** analýzy



(1) Rozřízneme obě nelineární proměnné na dvě skupiny respondentů a tím z nich vytvoříme dichotomní (= binární) proměnné.





## VÝSLEDKY binární **logistické** analýzy

Výzkumná otázka v logistické regresní analýze (= **kteří spolužáci získají nadprůměrnou míru empatie od spolužáků?**)

- Z předběžných analýz se ukázalo, že empatie dívek a chlapců může hrát odlišnou roli v získané empatii od spolužáků. Proto se v modelu zkoumala zvlášť u dívek a chlapců (viz interakce „**afekt.emp.\*pohlaví**“)

Tab. 11 Model binární logistické regrese

Nezávisle proměnné	B	SE	Wald	p	Exp(B)	95 % CI
Nagelkerke $R^2 = 0,30$						
Zkušenost s viktimizací (dich)	1,03	0,43	5,59	0,02	2,79	1,19-6,53
Pohlaví (d=0, ch=1)	6,53	2,52	6,70	0,01	686,89	4,9-96554
Afektivní emp. k druhým	1,07	0,55	3,77	0,05	2,91	0,99-8,55
Skupinový status (dich)	-2,88	0,79	13,37	0,00	0,06	0,01-0,26
Afekt. emp. k druh*pohlaví	-1,73	0,70	6,04	0,01	0,18	0,05-0,71
Absolutní člen	5,53	2,10	6,92	0,01	0,04	

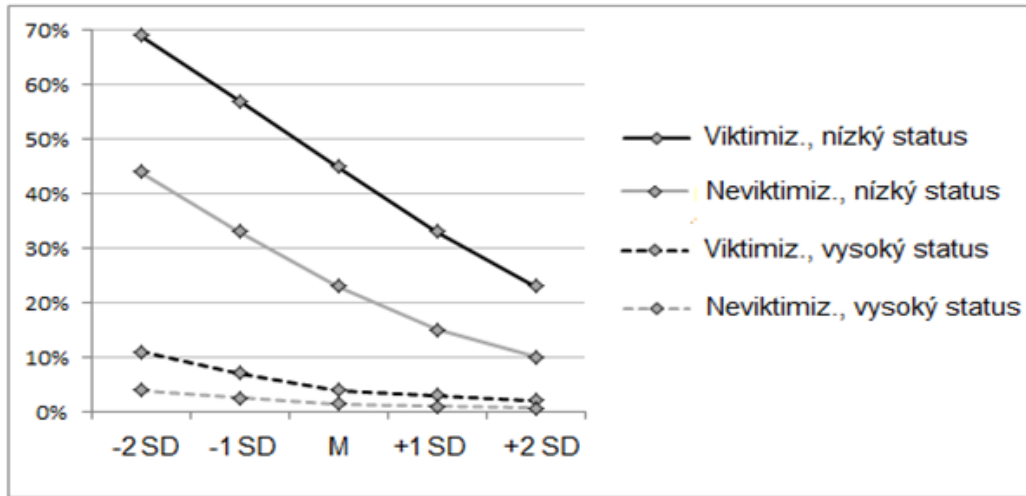
Pozn.: Závisle proměnná: nadprůměrná kognitivní empatie získaná od spolužáků (dichotomizovaná)

Cox-Snell  $R^2 = 0,19$

Stupně volnosti = 1

V binární logistické regresi se počítá **pravděpodobnost** či **šance**. Zde šance na to, že žák získá od spolužáků nadprůměrnou míru empatie.

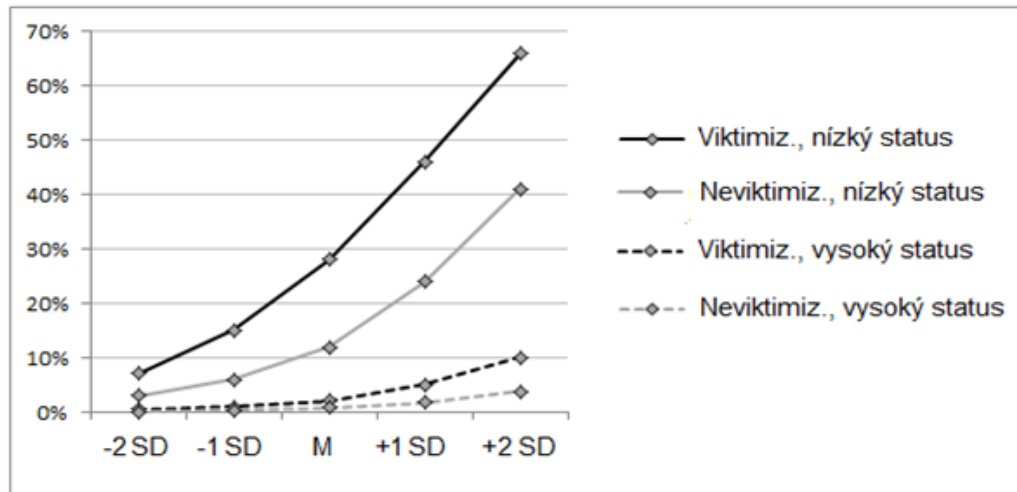
Obrázek 12 Pravděpodobnost **u chlapců**, že získají nadprůměrnou míru kognitivní empatie od spolužáků



ROZKRESLENÍ VÝSLEDKŮ  
binární **logistické** analýzy

Pozn.: Osa X znázorňuje zvyšující se afektivní empatii vůči druhým (hodnoty průměru [ $M = 3,35$ ] a směrodatné odchyly [ $SD = 0,75$ ] vycházejí z celého souboru)

Obrázek 11 Pravděpodobnost **u dívek**, že získají nadprůměrnou míru kognitivní empatie od spolužáků



Pozn.: Osa X znázorňuje zvyšující se afektivní empatii vůči druhým (hodnoty průměru [ $M = 3,35$ ] a směrodatné odchyly [ $SD = 0,75$ ] vycházejí z celého souboru)

# Děkuji Vám za pozornost

- přejdeme k Vaším dotazům, poznámkám a námětům:

