

# Regresní analýza

# Organizačně

- přednáška
  - Co je to regresní analýza?
  - Kdy se používá?
  - Na jaké otázky může nabídnout odpověď?
  - Základní principy
- Praktické procvičení
- Pokud něčemu neporozumíte, pak je zde možnost konzultace (po předchozí domluvě, takřka kdykoli)

# Použití

- TESTOVÁNÍ TEORIÍ !!!
- Zjištění vlivu nezávisle proměnné na závisle proměnnou
  - Při kontrole dalších možných faktorů
  - (predikce: jakou hodnotu bude mít závisle proměnná při určité kombinaci nezávisle proměnných)

# Příklady otázek ze závěrečných prací v ISu

- Jak závisí naděje na zvědavosti a přemítání?
  - H1: Respondenti dosahující vyšší míry zvědavosti vykazují vyšší míru naděje.
  - H2: Respondenti dosahující vyšší míry přemítání vykazují nižší míru naděje
- Jak závisí stupeň glomerulární filtrace na biochemických, demografických a antropometrických údajích pacientů?
  - Mezi nezávislé faktory asociované s nižší glomerulární filtrací patří vyšší hladina sérového kreatininu, vyšší věk, ženské pohlaví, jiný než Afroamerický etnický původ, vyšší koncentrace sérové urey a nižší koncentrace sérového albuminu
- Jak závisí rychlost plavání na stylu?

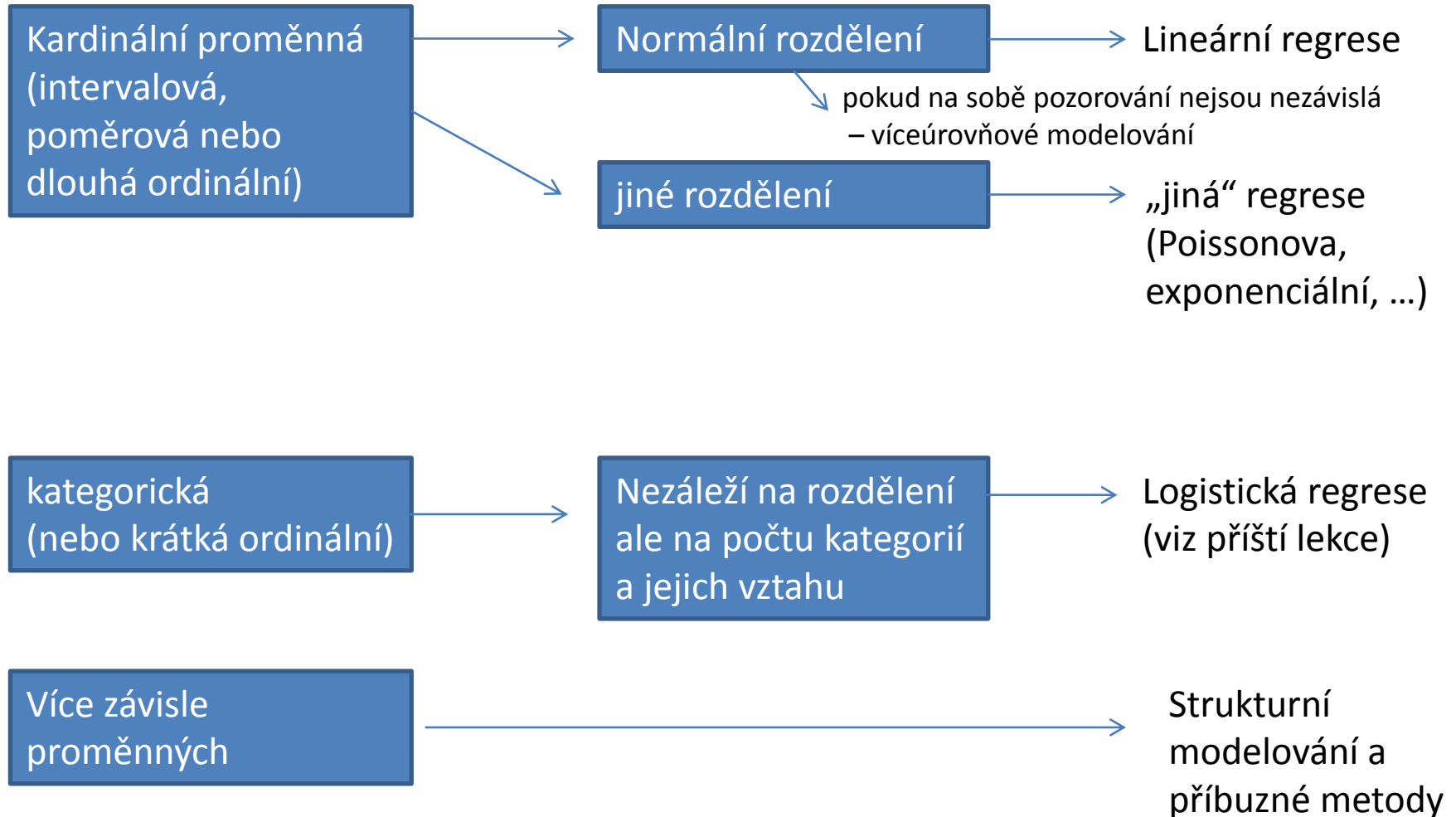
# Příklady otázek ze závěrečných politologických prací v ISu

- Desítky volebně-geografických prací
- Co ovlivňuje jednotu českých poslaneckých klubů?
  - H1a: Jednotnost hlasování je vyšší u vládních stran.
  - H1b: Jednotnost hlasování vládní strany je vyšší, čím těsnější je většina, kterou disponuje.
- Je míra korupce ovlivněna i používaným volebním systémem?
  - 1) Korupce roste s rostoucími volebními obvody v systémech s oteřevnými kandidátkami.
  - 2) Korupce klesá s rostoucími volebními obvody v systémech s uzavřenými kandidátkami.
- Co ovlivňuje (ne)účast poslanců na hlasování v Poslanecké sněmovně PČR?
  - účast na hlasování se bude zvyšovat s rostoucí pravděpodobností, že daný poslanec, či poslankyně bude pivotálním ... hlasem ...
  - poslanci ze vzdálenějších obvodů budou mít vyšší míru absencí při hlasováních ve Sněmovně než poslanci, kteří jsou přímo z Prahy, nebo blízkého okolí

# Podmínky

- Jedna závisle proměnná
  - + Jedna nebo více nezávisle proměnných
- Normálně rozdělená závisle proměnná
  - Nejde o úplně nezbytnou podmínku
  - Rozdělení a typ nezávisle proměnné může být jakékoli
- Několik dalších různě důležitých podmínek
  - Nezávislost pozorování
  - Předpoklad lineárního vztahu
  - Nezávislost nezávisle proměnných mezi sebou
  - Homogenní rozptyl reziduí

# Rozhodovací strom

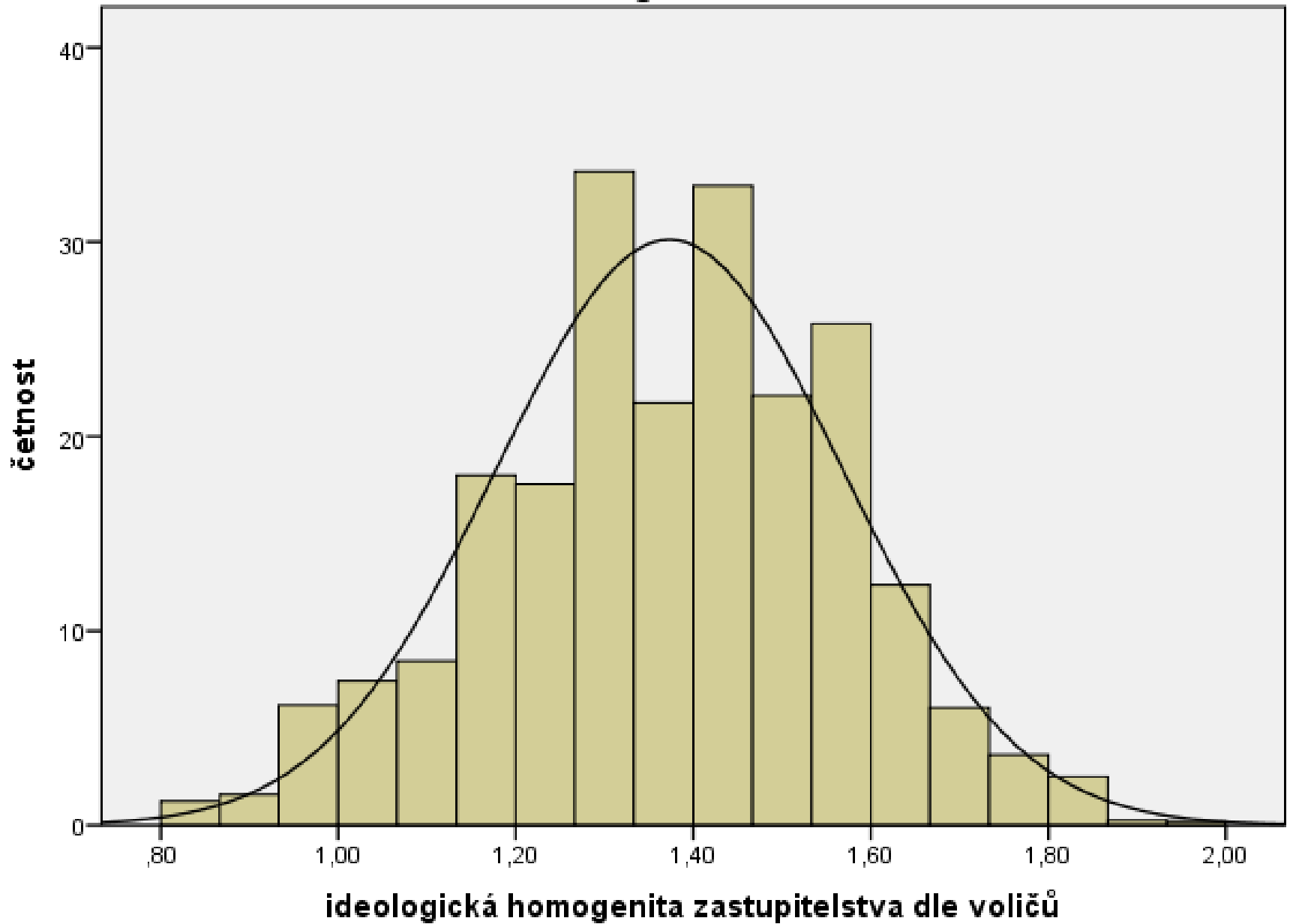




	Stata	SAS	SPSS	Mplus	R
<b>Regression Models</b>					
Robust Regression	<a href="#">Stata</a>	<a href="#">SAS</a>			<a href="#">R</a>
<b>Models for Binary and Categorical Outcomes</b>					
Logistic Regression	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>	<a href="#">Mplus</a>	<a href="#">R</a>
Exact Logistic Regression	<a href="#">Stata</a>	<a href="#">SAS</a>			<a href="#">R</a>
Multinomial Logistic Regression	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>	<a href="#">Mplus</a>	<a href="#">R</a>
Ordinal Logistic Regression	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>	<a href="#">Mplus</a>	<a href="#">R</a>
Probit Regression	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>	<a href="#">Mplus</a>	<a href="#">R</a>
<b>Count Models</b>					
Poisson Regression	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>	<a href="#">Mplus</a>	<a href="#">R</a>
Negative Binomial Regression	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>	<a href="#">Mplus</a>	<a href="#">R</a>
Zero-inflated Poisson Regression	<a href="#">Stata</a>	<a href="#">SAS</a>		<a href="#">Mplus</a>	<a href="#">R</a>
Zero-inflated Negative Binomial Regression	<a href="#">Stata</a>	<a href="#">SAS</a>		<a href="#">Mplus</a>	<a href="#">R</a>
Zero-truncated Poisson	<a href="#">Stata</a>	<a href="#">SAS</a>			<a href="#">R</a>
Zero-truncated Negative Binomial	<a href="#">Stata</a>	<a href="#">SAS</a>		<a href="#">Mplus</a>	<a href="#">R</a>
<b>Censored and Truncated Regression</b>					
Tobit Regression	<a href="#">Stata</a>	<a href="#">SAS</a>		<a href="#">Mplus</a>	<a href="#">R</a>
Truncated Regression	<a href="#">Stata</a>	<a href="#">SAS</a>			<a href="#">R</a>
Interval Regression	<a href="#">Stata</a>	<a href="#">SAS</a>			<a href="#">R</a>
<b>Multivariate Analysis</b>					
One-way MANOVA	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>		
Discriminant Function Analysis	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>		
Canonical Correlation Analysis	<a href="#">Stata</a>	<a href="#">SAS</a>	<a href="#">SPSS</a>		<a href="#">R</a>
Multivariate Multiple Regression	<a href="#">Stata</a>	<a href="#">SAS</a>		<a href="#">Mplus</a>	
<b>Mixed Effects Models</b>					
Generalized Linear Mixed Models	<a href="#">Introduction to GLMMs</a>				
Mixed Effects Logistic Regression	<a href="#">Stata</a>				<a href="#">R</a>
<b>Other</b>					

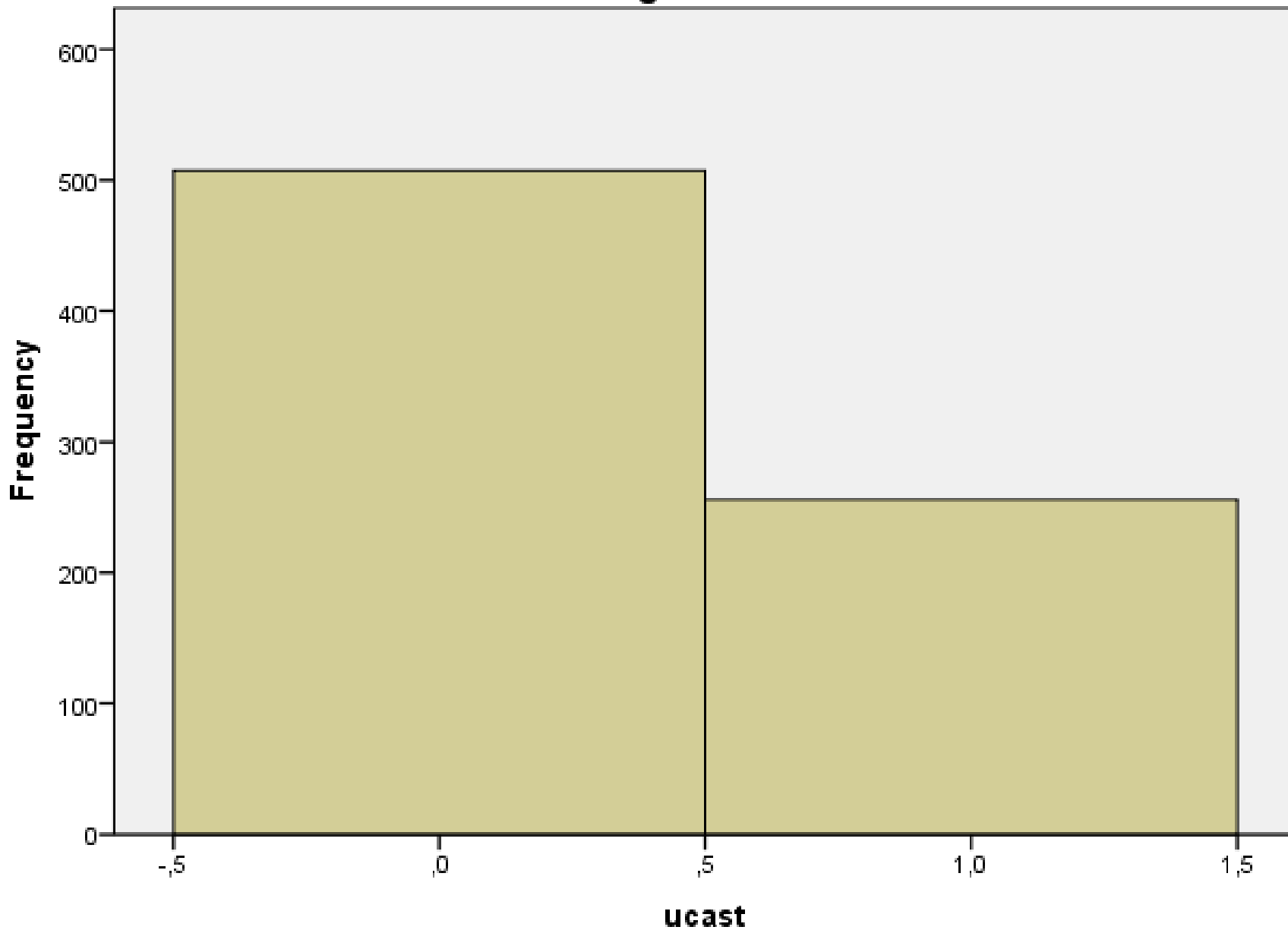


# Histogram 1



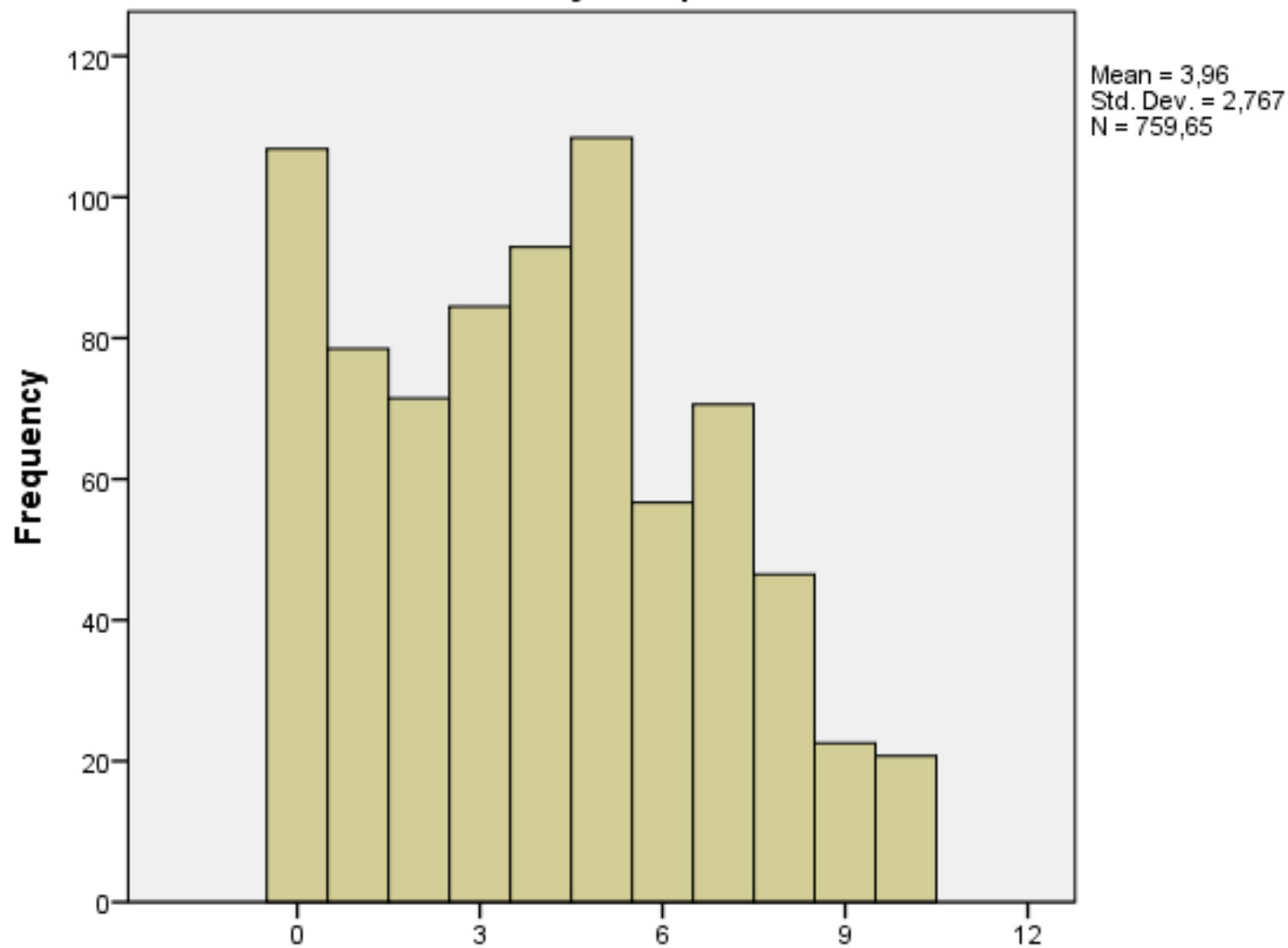
Cases weighted by vaha

# Histogram 2



Cases weighted by vaha

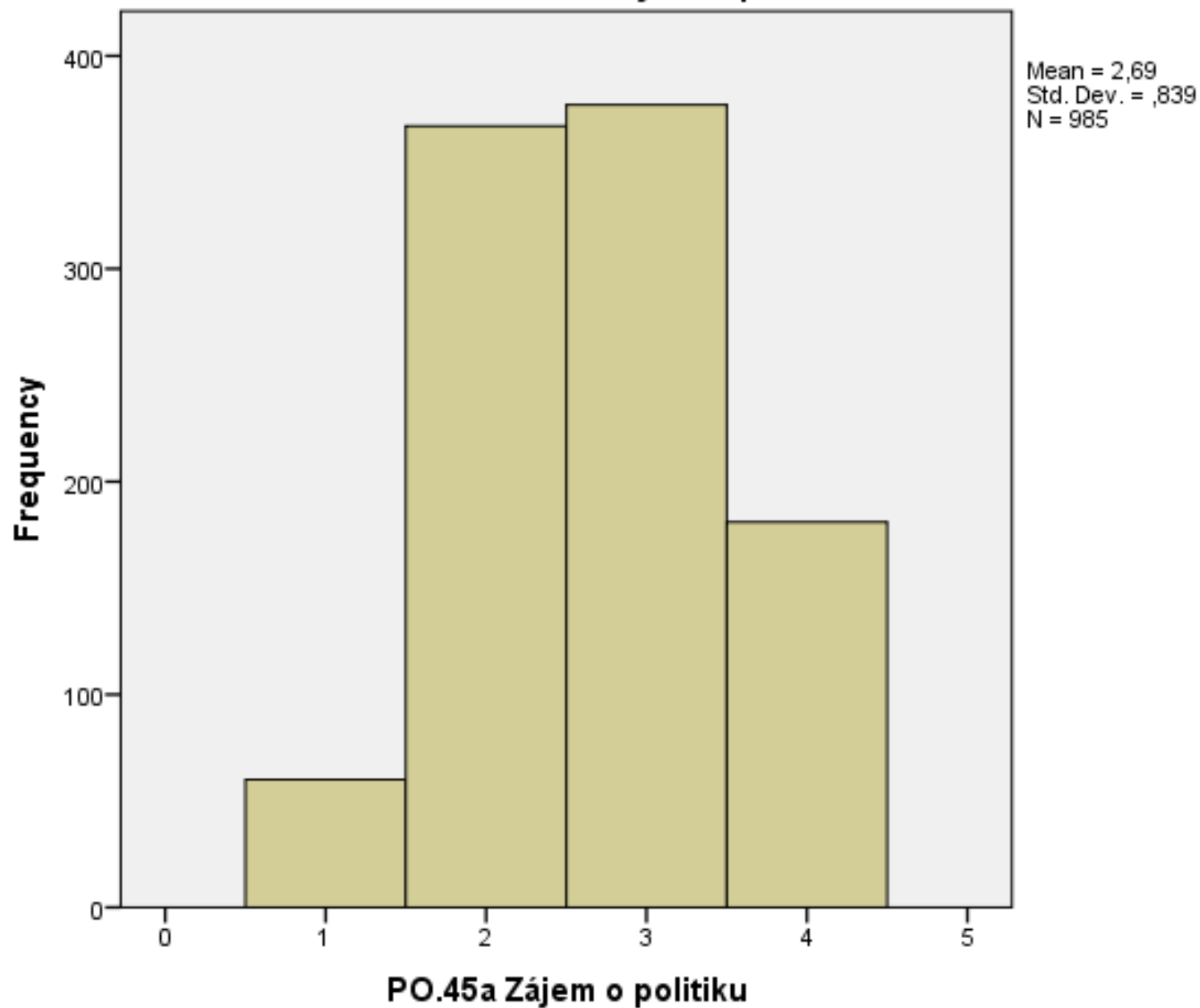
## zájem o politiku: EU



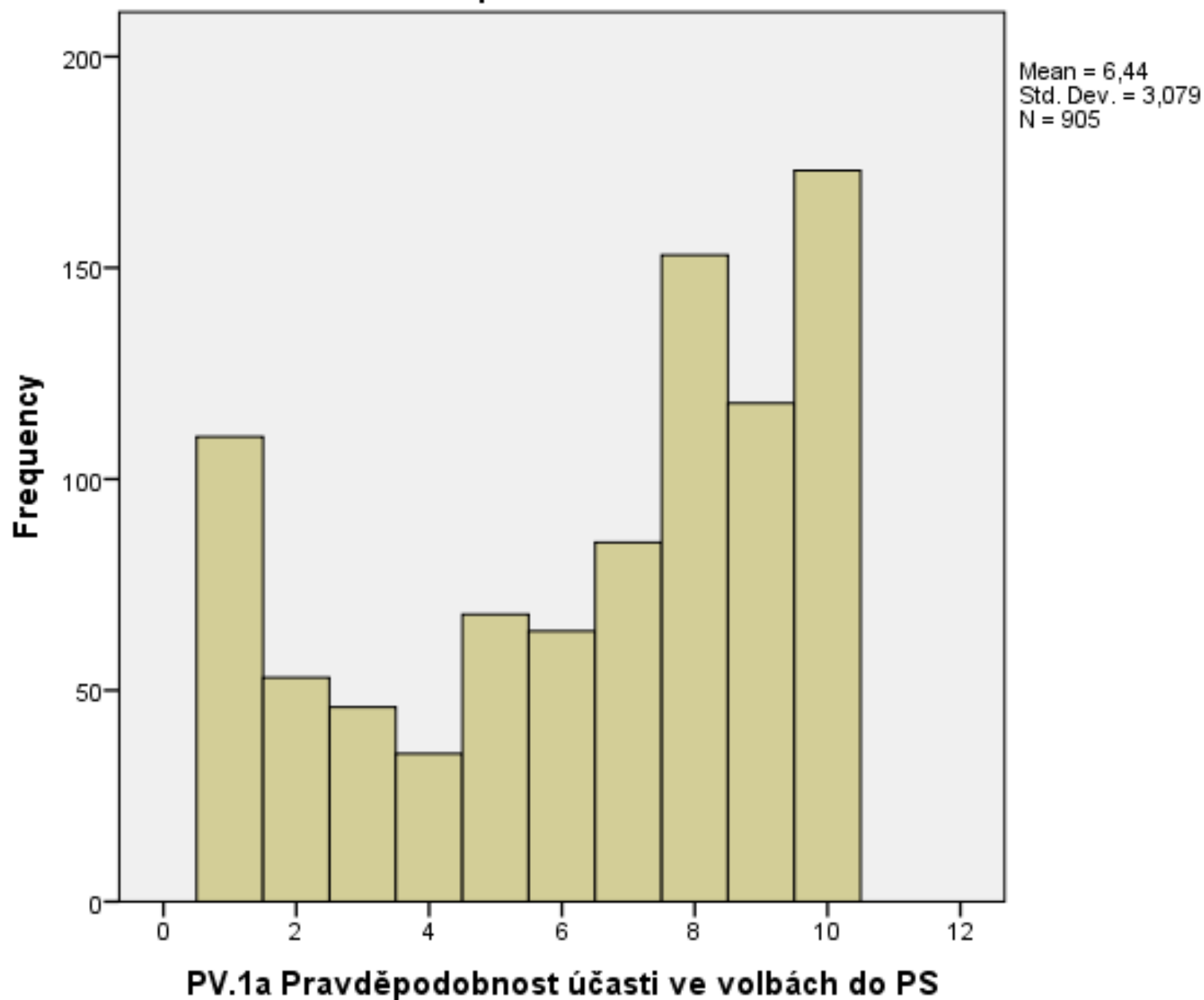
zájem o politiku: EU

Cases weighted by vaha

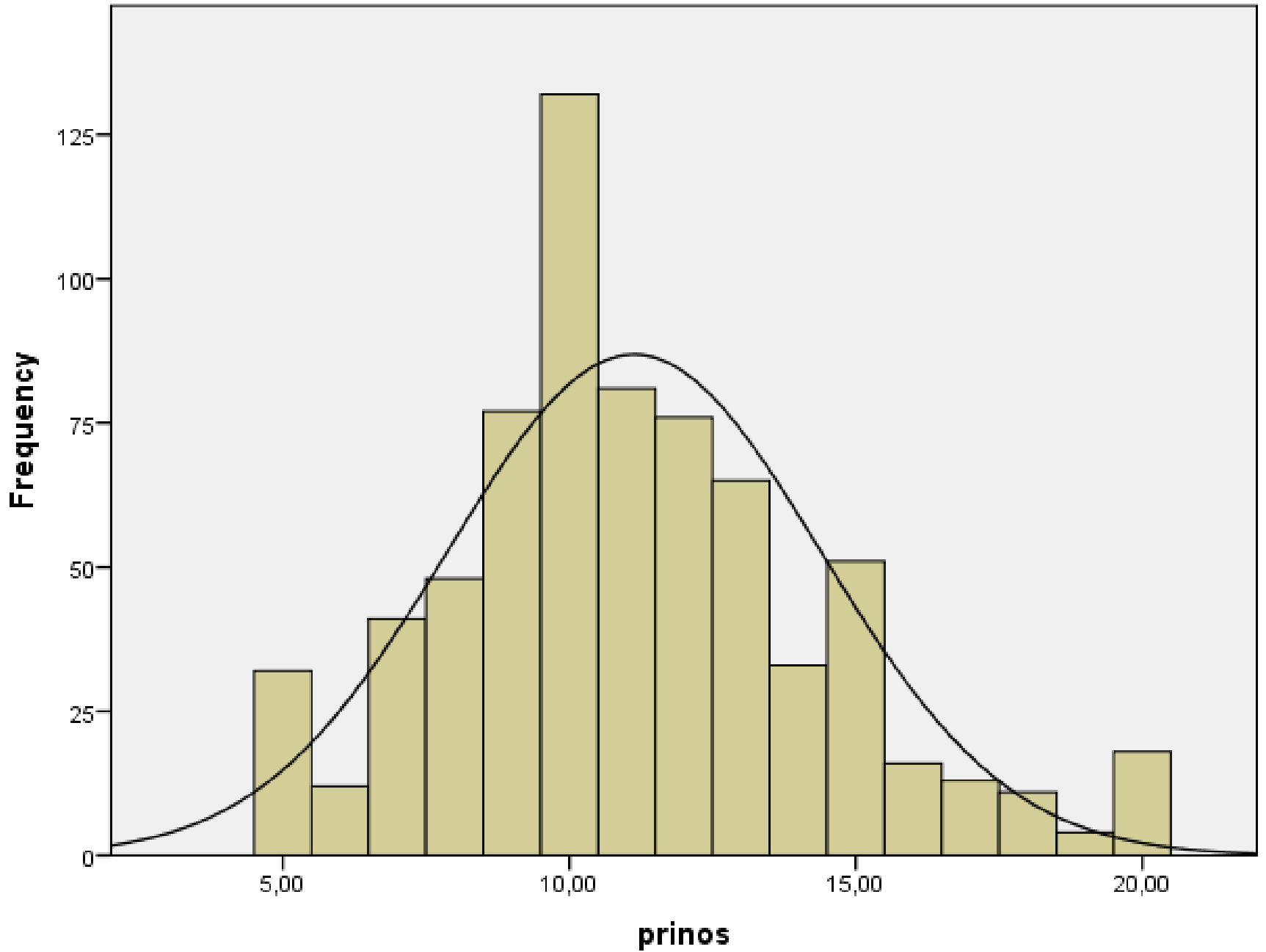
### PO.45a Zájem o politiku



## PV.1a Pravděpodobnost účasti ve volbách do PS



# Histogram



# Co regrese dělá

- Odhad parametrů přímky při 1 nezávisle proměnné (roviny při 2 np či nadroviny při více np)
- Parametry: sklon (pro každou proměnnou) a konstanta (jedna pro celý model)
- Parametry popisují vztah mezi nezávisle a závisle proměnnou
- Parametry tvoří rovnici přímky
- Hodnota závisle proměnné = konstanta + sklon \* hodnota nezávisle proměnné
- $y = a + b * x + e$
- $y = a + b_1 * x + b_2 * x + b_3 * x + \dots + e$

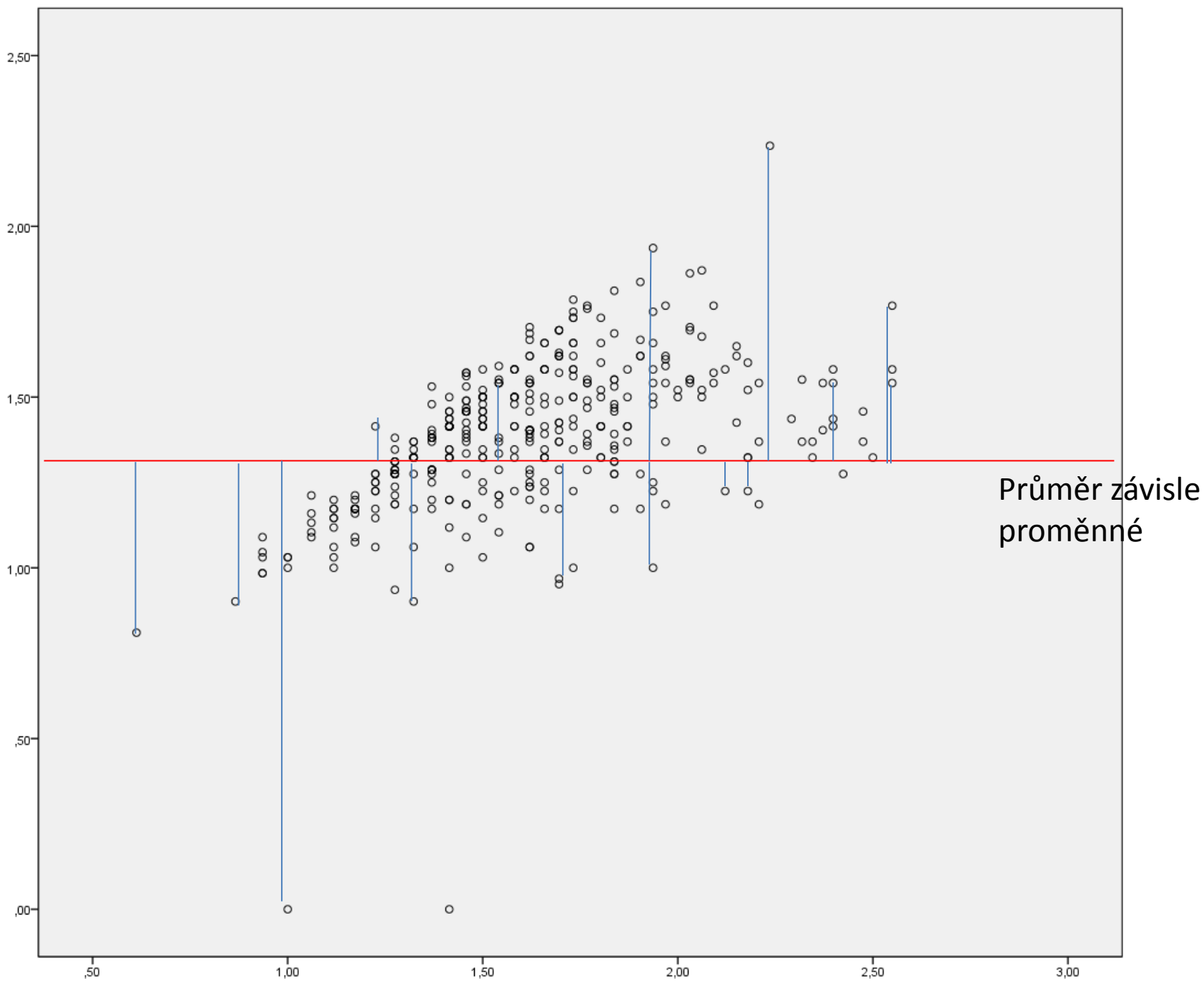
# Co nám výpočet poskytne?

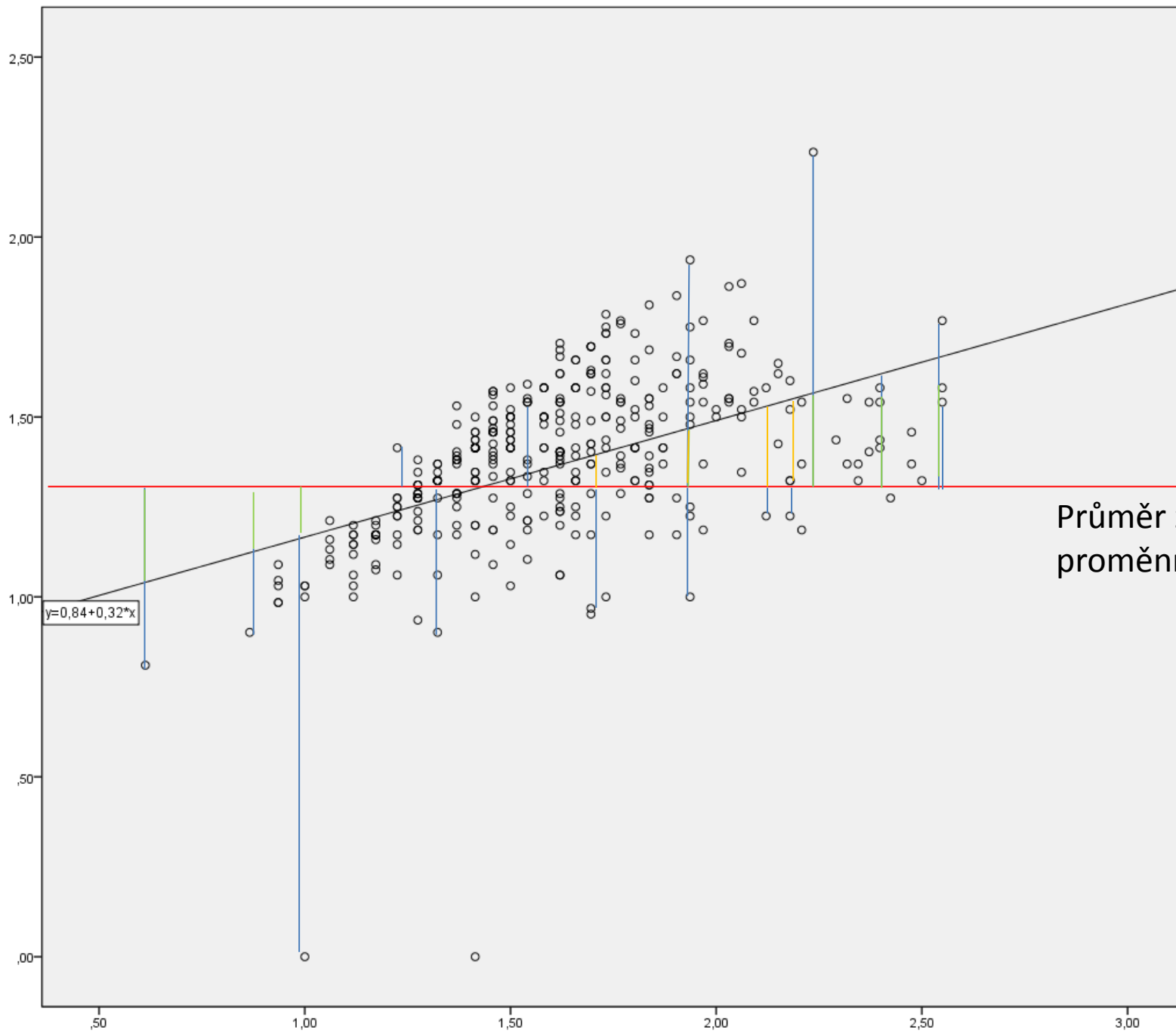
- R-square ( česky index determinace)
  - Ukazuje jak dobře model sedí na data
- Parametry
  - Unstandardized beta (nestandardizovaný beta koeficient)
  - Constant (konstanta)
- Hodnoty signifikance



# Co je to R-square?

- Ukazuje, kolik procent rozptylu závisle proměnné je vysvětleno přidáním nezávisle proměnných
- Původní rozptyl je vypočten jako suma kvadratických odchylek mezi průměrem a jednotlivými hodnotami závisle proměnné
- „nový“ rozptyl je vypočten jako suma odchylek od regresní přímky/roviny
- Rozdíl mezi původním a novým rozptylem vydělený původní variabilitou = R-square
- Čím víc proměnných, tím nižší R-square
  - Řešeno pomocí adjusted R-square





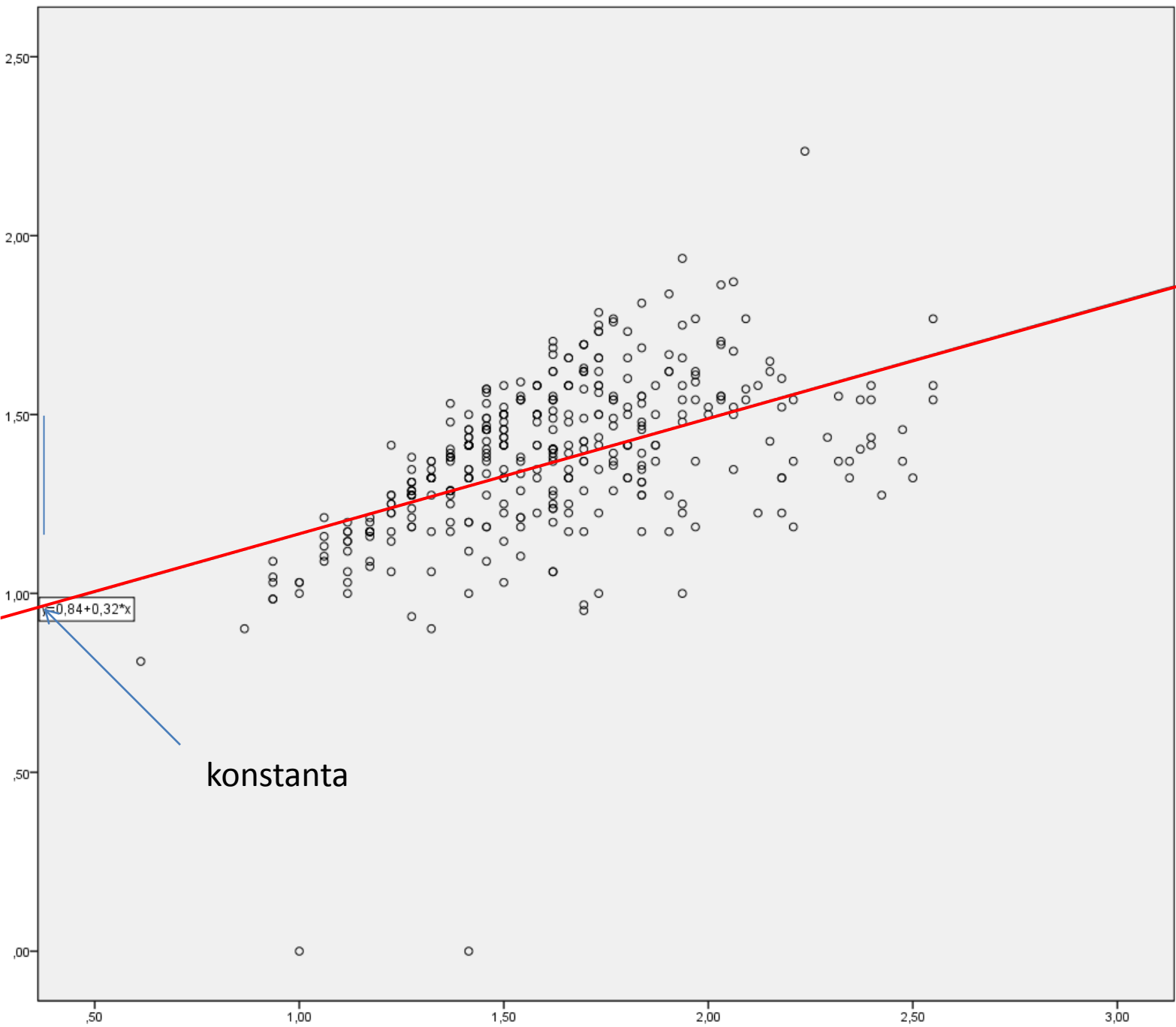
Průměr závisle  
proměnné

$y = 0,84 + 0,32 \cdot x$

# Konstanta

- Jaká je očekávaná hodnota nezávisle proměnné, pokud jsou hodnoty všech nezávisle proměnných 0
- Pro smysluplnou interpretaci je často potřeba rekódovat proměnné
  - Každý má nějaký věk, pohlaví, výšku, váhu, ...

Y



konstanta

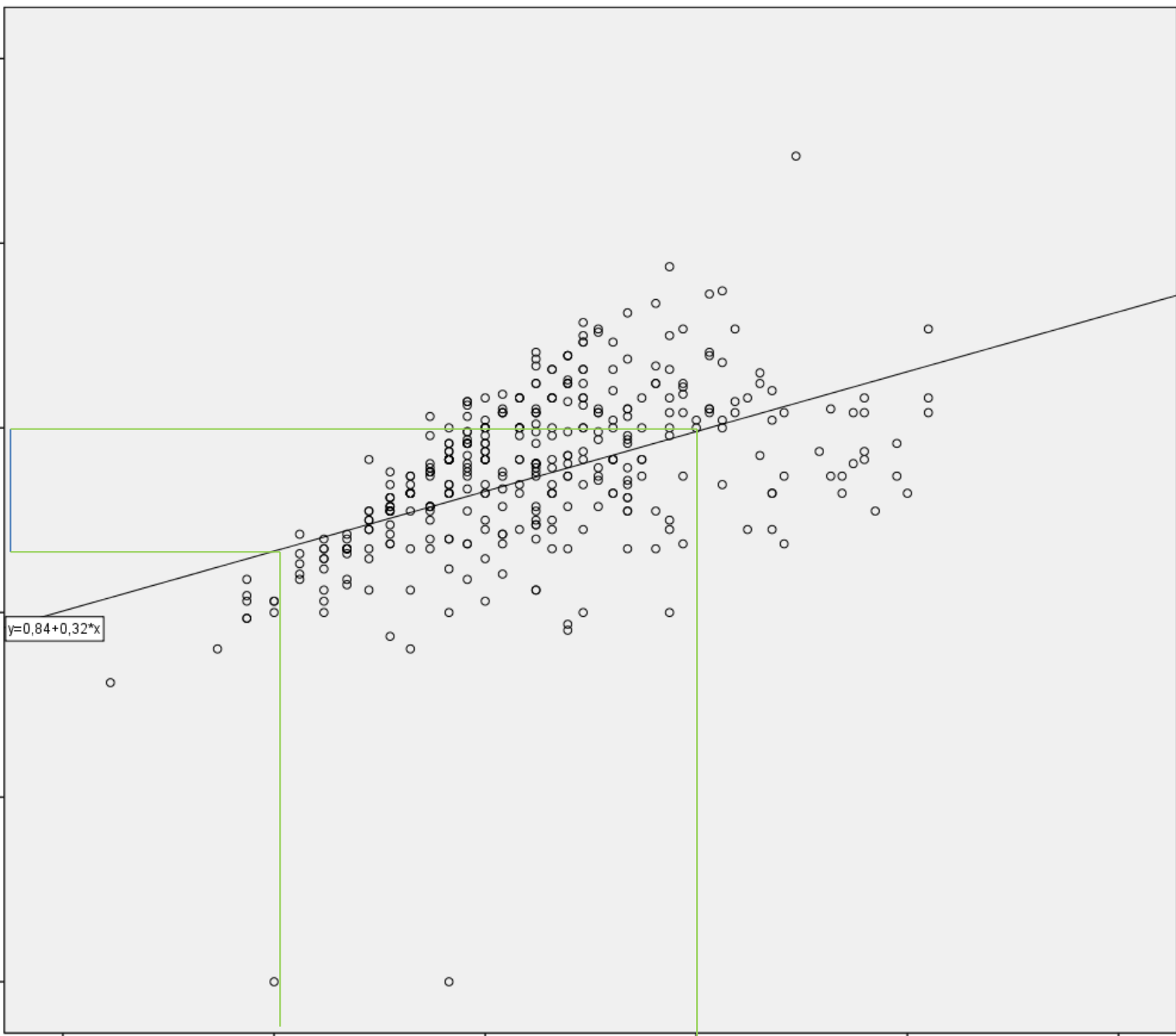
X

# Nestandardizovaný Beta koeficient

- „o kolik se změní hodnota závisle proměnné, pokud se hodnota nezávisle proměnné změní o jednotku“
- Různé proměnné se mohou změnit o různý počet jednotek
  - Pro srovnání síly proměnných v modelu – standardizovaný koeficient beta ( jakou změnu v počtu směrodatných odchylek závisle proměnné způsobí změna o směrodatnou odchylku nezávisle proměnné)

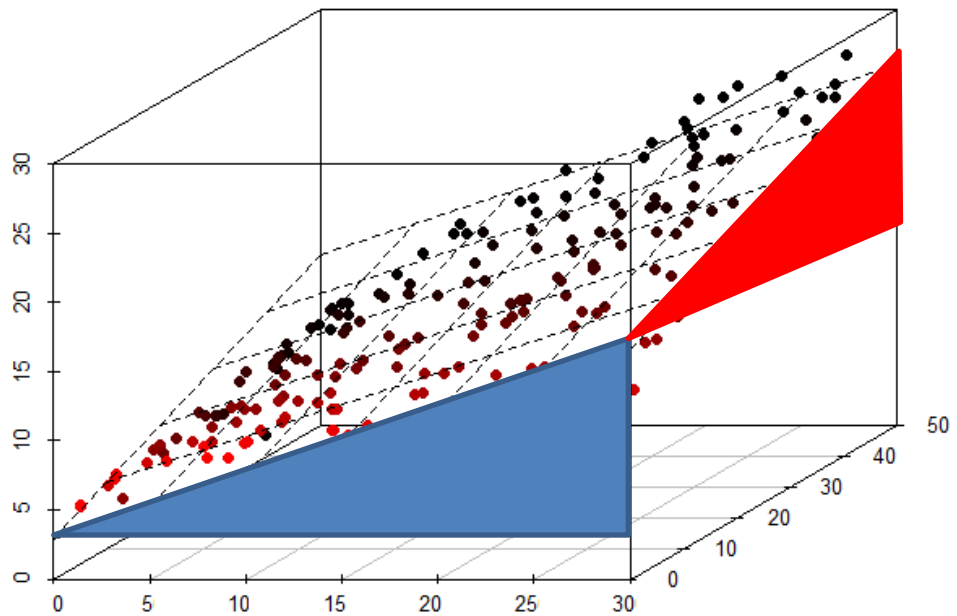
Y

2,50  
2,00  
1,50  
1,00  
,50  
,00



$y = 0,84 + 0,32 * x$

X





# Příklad

- Téma: Názory na zasahování státu do ekonomiky
- Popis problému:
  - Občané mají různé názory na to, zda a jak by měl stát zasahovat do hospodářství
- Otázka: Co způsobuje rozdílné názory na zásahy státu do ekonomiky mezi občany?

# Postup

- Nadefinování modelu pomocí hypotéz vycházejících z teorie
- Sestavení datasetu obsahujícího závisle a nezávisle proměnné dle specifikace
- Zkontrolování normality závisle proměnné
- Zkontrolování vlastností nezávisle proměnných

# Teorie

- Politické hodnoty
- Hodnoty jsou preferovanými stavy věcí (svoboda x sociální spravedlnost)
- Hodnoty se utváří v průběhu socializace – role věku
- Hodnoty jsou ovlivněny aktuální situací jedince (adaptace) - role příjmu
- Role vzdělání a třídy

# Hypotézy

- H1: starší voliči budou preferovat vyšší míru zasahování státu do ekonomiky
- H2: s rostoucím příjmem poroste preference vyšší ekonomické svobody.
- H3: s vyšším vzděláním poroste preference vyšší ekonomické svobody
- H3X: s vyšším vzděláním poroste preference vyšší míry zasahování státu do ekonomiky
- H4: lidé se zkušeností s nezaměstnaností budou více preferovat zásahy do ekonomiky než lidé bez takové zkušenosti
  
- H0 proměnná nemá vliv

# Data

- Naše společnost CVVM
  - Dataset s baterií otázek na ekonomickou levici/pravici
  - dataset obsahuje i proměnné věk, vzdělání, příjem a zkušenost s nezaměstnaností

# Proměnné

- Závisle proměnná: Míra zasahování státu do ekonomiky
  - Vytvořeno jako faktorové skóre na základě proměnných
  - Hodnoty 0 – 10 (0 – zasahování, 10 – svoboda)
  - Ke kterému z každé dvojice následujících výroků byste se spíše přiklonil?
  - Rozvoj hospodářství má být ponechán vlastnímu vývoji/má být usměrňován státem
  - Stát má zaručit, aby ten, kdo chce pracovat, dostal práci/  
Kdo chce pracovat, musí se o získání práce postarat sám
  - Velkým hospodářským podnikům má stát umožnit co největší samostatnost/ Na velké hospodářské podniky má stát co nejvíce dohlížet
  - Velikost soukromého vlastnictví by nijak být omezována neměla/by nějakým způsobem být omezována měla.

# Nezávisle proměnné

- Nezávisle / kontrolní proměnné
- Příjem: čistý příjem domácnosti
- Subjektivní hodnocení příjmu (dichotomická proměnná)
- Věk
- Nespokojenost s vnějšími podmínkami: součet proměnných ptajících se na hodnocení ekonomické a politické situace (od 0 do 10)
- Vzdělání: kategorická proměnná rekódovaná na dummy proměnné
  - ZŠ vzdělání referenční kategorií
- Nezaměstnanost: kategorická proměnná rekódovaná na dummy proměnné
  - Bez zkušenosti s nezaměstnaností jako referenční kategorie

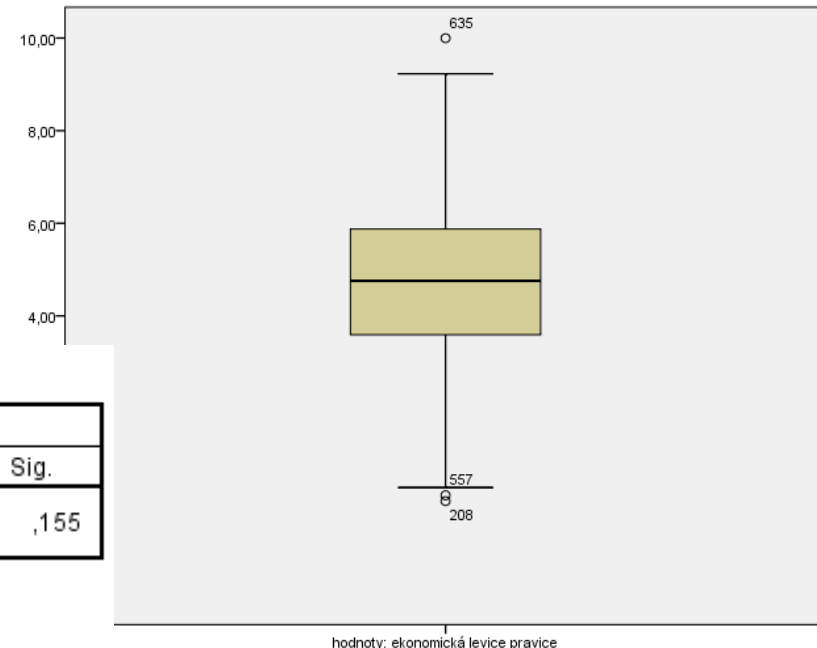
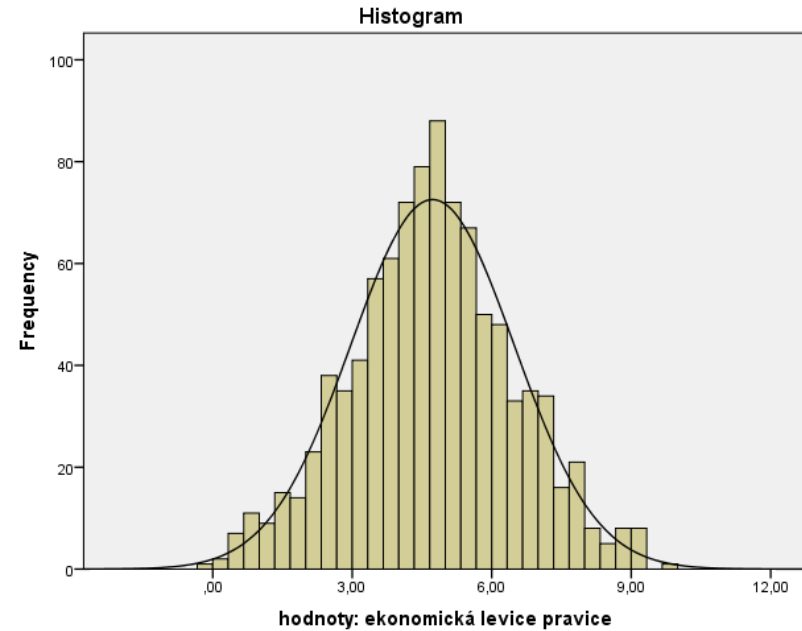
# Normalita závisle proměnné

- Jinakost rozdělení
  - ovlivňuje především hodnoty signifikance
  - Zkresluje odhady parametrů
- Prvně vizuální zhodnocení pomocí histogramu
- Testy
  - K-S a S-W
  - Ve velkých souborech lze brát s rezervou
  - Šikmost a strmost není větší než  $3 \times SE$



# Test normality závisle proměnné

- Histogram
  - Analyze- descriptive stat- frequencies – plots
- Kolmogorův-Smirnovův test
  - Analyze – descriptive stat – explore – plots – normality plots with tests



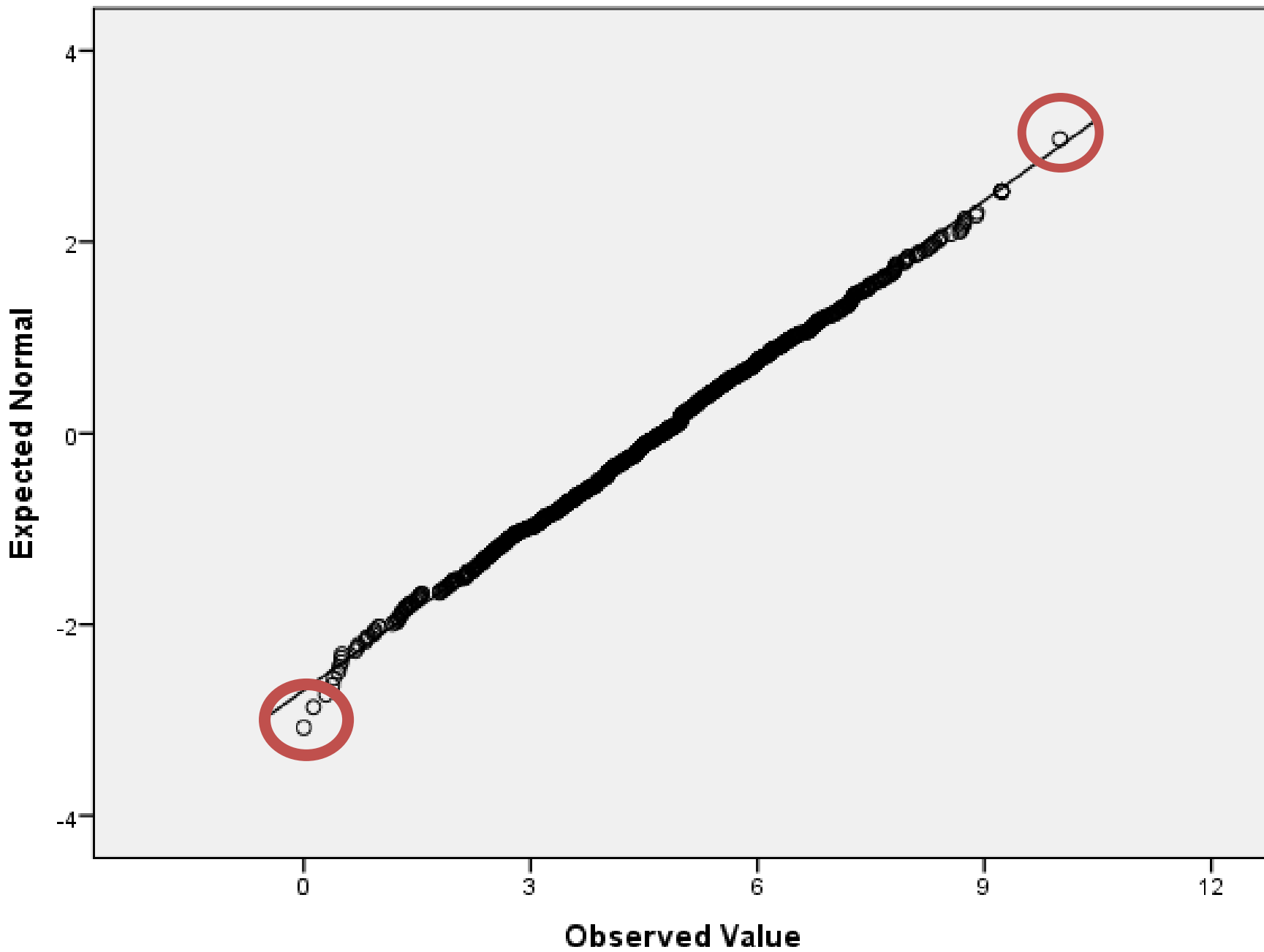
**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
hodnoty: ekonomická levice pravece	,021	959	,200 <sup>*</sup>	,998	959	,155

\*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Normal Q-Q Plot of hodnoty: ekonomická levice pravice



# Odebrání outlierů

- Outliery je možné z analýzy vynechat
- Jde o přípustnou manipulaci s daty
- Nutné reportovat!!!
- Vhodné také ukázat rozdíl ve výsledcích analýzy před a po odstranění outlierů

# Další postup

- Pokud je závisle proměnná v pořádku
  - Rekódování nezávisle proměnných
  - Kontrola multikolinearity nezávisle proměnných
    - Nezávisle proměnné by mezi sebou neměly příliš souviset
    - První kontrola pomocí korelačního koeficientu
    - Další kontrola přímo v modelu
  - Výpočet

# Průzkum souvislosti mezi proměnnými

- Crostab
- Existuje poměrně silný vztah mezi subjektivní chudobou a zkušeností s nezaměstnaností
- Rekódování kombinace proměnných

# Kontrola multikolinearity

- Analyze – correlate - bivariate

Correlations

		IDE.2 Věk	prijem	nespokojeno st s vnějšími podmínkami	IDE.10a Osobní čistý měsíční příjem
IDE.2 Věk	Pearson Correlation	1	-,330**	,155**	,163**
	Sig. (2-tailed)		,000	,000	,000
	N	1043	651	989	764
prijem	Pearson Correlation	-,330**	1	-,259**	,545**
	Sig. (2-tailed)	,000		,000	,000
	N	651	652	634	639
nespokojenost s vnějšími podmínkami	Pearson Correlation	,155**	-,259**	1	-,103**
	Sig. (2-tailed)	,000	,000		,006
	N	989	634	991	718
IDE.10a Osobní čistý měsíční příjem	Pearson Correlation	,163**	,545**	-,103**	1
	Sig. (2-tailed)	,000	,000	,006	
	N	764	639	718	765

\*\* . Correlation is significant at the 0.01 level (2-tailed).

# Naklikání modelu

- Analyze – regression – linear
- Dependent: lp\_ekonom
- Independent: vek, nespokojenost, prijem, chudi\_subj, učeň, sš, vš, zkus\_nezam, zajem, muž, mesto
- Statistics: colinearity diagnostics, casewise diagnostics >2,5
- Plots: Y:\*ZRESID, X:\*ZPRED

OK

# Interpretace $R^2$ a adj. $R^2$

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

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,306 <sup>a</sup>	,094	,077	1,69542	1,605

a. Predictors: (Constant), příjem, vek, ...

b. Dependent Variable: hodnoty: ekonomická levice pralice

- Model vysvětluje 9,4 % variability závisle proměnné

**ANOVA<sup>a</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	172,310	11	15,665	5,450	,000 <sup>b</sup>
	Residual	1664,315	579	2,874		
	Total	1836,625	590			

a. Dependent Variable: hodnoty: ekonomická levice pralice

b. Predictors: (Constant), příjem, vek, ...

- Model je statisticky významný (tj. můžeme jeho výstupy zobecnit na populaci)



# Interpretace $R^2$ a adj. $R^2$

- **neukazuje**, nakolik jsou výsledky platné v celém souboru,
- **neukazuje**, pro jaké procento voličů vztah platí
- ukazuje jak moc model vysvětluje rozptyl v závisle proměnné.
- Jak dobře model popisuje realitu (zaznamenanou v datech)
- Když je model nesignifikantní (tj. žádná z proměnných nepřispívá k vysvětlení rozptylu), tak použité proměnné nejsou vhodné,
  - nikoli, že k analýze proměnné není regrese použitelná
    - To závisí na naplnění předpokladů

# Interpretace konstanty

- Nesmyslná, protože nikdo ve vzorku nemá věk 0
- Proto proměnnou věk rekódujeme
  - Odečteme 15
- V novém modelu je konstantu možné interpretovat:
- hodnota závisle proměnné očekávaná pro nejmladší občanky, spokojené s podmínkami, bez příjmu, ale subjektivně bohaté, se zš vzděláním, bez zájmu o politiku, a zkušenosti s nezaměstnaností žijící ve vsi (= 5,1)

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	C
		B	Std. Error	Beta			
1	(Constant)	5,135	,426		12,064	,000	
	vek	-,018	,005	-,180	-4,057	,000	
	nespokojenost s vnějšími podmínkami	-,029	,044	-,028	-,651	,515	
	příjem	,010	,005	,090	1,932	,054	
	subjektivně chudá domácnost	-,367	,174	-,094	-2,102	,036	
	učňovské vzdělání nebo sš bez maturity	,086	,210	,024	,412	,681	
	sš s maturitou	,181	,223	,048	,815	,415	
	vysokoškolské vzdělání	,272	,264	,055	1,030	,303	
	zkušenost s nezaměstnaností	-,273	,186	-,062	-1,470	,142	
	zajímá se o politiku	-,053	,149	-,015	-,353	,724	
	muž	,111	,145	,031	,768	,443	
	obec nad 5000 obyvatel	,069	,151	,018	,458	,647	

a. Dependent Variable: hodnoty: ekonomická levice pravice

# Interpretace nestandardizovaného beta koeficientu

- 2 situace
- Dummy proměnné x kardinální proměnné
- Interpretace efektu dummy proměnné:
  - nestandardizovaný koeficient ukazuje rozdíl dané kategorie oproti referenční kategorii
- Interpretace efektu kardinální proměnné
  - Při změně nezávisle proměnné o jednotku se hodnota závisle proměnné změní o hodnotu nestandardizovaného koeficient

# Interpretace efektu dummy proměnné

Coefficients<sup>a</sup>

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Total
	B	Std. Error	Beta			
1	(Constant)	5,135	,426			
	vek	-,018	,005	-,180	-,4057	,000
	nespokojenost s vnějšími podmínkami	-,029	,044	-,028	-,651	,515
	příjem	,010	,005	,090	1,932	,054
	subjektivně chudá domácnost	-,367	,174	-,094	-2,102	,036
	učňovské vdělání nebo sš bez maturity	,086	,210	,024	,412	,681
	sš s maturitou	,181	,223	,048	,815	,415
	vysokoškolské vzdělání	,272	,264	,055	1,030	,303
	zkušenost s nezaměstnaností	-,273	,186	-,062	-1,470	,142
	zajímá se o politiku	-,053	,149	-,015	-,353	,724
	muž	,111	,145	,031	,768	,443
	obec nad 5000 obyvatel	,069	,151	,018	,458	,647

a. Dependent Variable: hodnoty: ekonomická levice pravice

# Interpretace efektu dummy proměnné

- Subjektivně chudý občan preferuje zásahy do ekonomiky více než subjektivně bohatý občan volič (pokud jsou ostatní sledované charakteristiky stejné) a to o 0,36 bodu
- Nebo též
- Pokud je vše ostatní shodné, pak rozdíl na škále ekonomických hodnot mezi bohatým a chudým občanem je 0,36 bodu . Chudý občan více preferuje zásahy do ekonomiky .

# Interpretace efektu kardinální

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	C
		B	Std. Error	Beta			
1	(Constant)	5,135	,426		12,064	,000	
	vek	-,018	,005	-,180	-4,057	,000	
	nespokojenost s vnějšími podmínkami	-,029	,044	-,028	-,651	,515	
	prijem	,010	,005	,090	1,932	,054	
	subjektivně chudá domácnost	-,367	,174	-,094	-2,102	,036	
	učňovské vzdělání nebo sš bez maturity	,086	,210	,024	,412	,681	
	sš s maturitou	,181	,223	,048	,815	,415	
	vysokoškolské vzdělání	,272	,264	,055	1,030	,303	
	zkušenost s nezaměstnaností	-,273	,186	-,062	-1,470	,142	
	zajímá se o politiku	-,053	,149	-,015	-,353	,724	
	muž	,111	,145	,031	,768	,443	
	obec nad 5000 obyvatel	,069	,151	,018	,458	,647	

a. Dependent Variable: hodnoty: ekonomická levice pravice

# Interpretace efektu kardinální proměnné

- Pokud má občan A o 1 000 Kč vyšší příjem než volič B a vše ostatní je shodné, měl by volič A o preferovat o 0,01 svobodnější ekonomiku
- Nebo též
- Pokud příjem vzroste o 1000 Kč a vše ostatní zůstane shodné, pak preference ekonomické svobody vzroste o 0,01 bodu
- Lze násobit
  - Pokud příjem vzroste o 10 000 Kč, pak preference ekonomické svobody vzroste o 0,1 bodu
  - Pokud příjem vzroste o 100 000 Kč, pak preference ekonomické svobody vzroste o 1 bod



### Coefficients<sup>a</sup>

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1	(Constant)	5,135	,426		
	vek	-,018	,005	-,180	-4,057
	nespokojenost s vnějšími podmínkami	-,029	,044	-,028	-,651
	příjem	,010	,005	,090	1,932
	subjektivně chudá domácnost	-,367	,174	-,094	-2,102
	učňovské vzdělání nebo sš bez maturity	,086	,210	,024	,412
	sš s maturitou	,181	,223	,048	,815
	vysokoškolské vzdělání	,272	,264	,055	1,030
	zkušenost s nezaměstnaností	-,273	,186	-,062	-1,470
	zajímá se o politiku	-,053	,149	-,015	-,353
	muž	,111	,145	,031	,768
	obec nad 5000 obyvatel	,069	,151	,018	,458

a. Dependent Variable: hodnoty: ekonomická levice pravice

# Hodnocení signifikance

- Zobecňování výsledků na populaci
- Obvyklá hranice sig.  $< 0,05$
- Potom považujeme efekt za signifikantní na hladině významnosti 95 %
- Nic nám nebrání zvolit si jinou hladinu významnosti (např. 90%, 99% nebo 99,99%)
- S nižší hladinou roste riziko, že budeme za platný považovat i efekt který v populaci neplatí
- S vyšší hladinou vyšší riziko že budeme za neplatný považovat i efekt, který v populaci platí

# Následná kontrola

- Outlieři
- Homogenita rozptylu reziduí (homoskedascita)
- multikolinearita

# Honocení multikolinearity

- VIF
- Arbitární hranice: 5
- A zároveň podobné hodnoty v dimenzích
  
- Proměnné levice a pravice
  - V pořádku, neboť se jedná o dummy proměnné vytvořené z jedné kategorické proměnné

Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	5,135	,426		12,064	,000		
	vek	-,018	,005	-,180	-4,057	,000	,793	1,262
	nespokojenost s vnějšími podmínkami	-,029	,044	-,028	-,651	,515	,859	1,161
	příjem	,010	,005	,090	1,932	,054	,723	
	subjektivně chudá domácnost	-,367	,174	-,094	-2,102	,036	,790	
	učňovské vzdělání nebo sš bez maturity	,086	,210	,024	,412	,681	,471	
	sš s maturitou	,181	,223	,048	,815	,415	,457	
	vysokoškolské vzdělání	,272	,264	,055	1,030	,303	,555	
	zkušenost s nezaměstnaností	-,273	,186	-,062	-1,470	,142	,890	
	zajímá se o politiku	-,053	,149	-,015	-,353	,724	,879	1,138
	muž	,111	,145	,031	,768	,443	,931	1,074
	obec nad 5000 obyvatel	,069	,151	,018	,458	,647	,960	1,042

a. Dependent Variable: hodnoty: ekonomická levice pravice

Collinearity Diagnostics<sup>a</sup>

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions										
				(Constant)	vek	nespokojeno st s vnějšími podmínkami	příjem	subjektivně chudá domácnost	učňovské vzdělání nebo sš bez maturity	sš s maturitou	vysokoškolské vzdělání	zkušenost s nezaměstnaností	zajímá se o politiku	
1	1	6,681	1,000	,00	,00	,00	,00	,00	,00	,00	,00	,00	,00	,01
	2	1,189	2,370	,00	,00	,00	,01	,07	,05	,01	,12	,12	,01	,01
	3	1,010	2,573	,00	,00	,00	,00	,00	,04	,18	,13	,01	,00	,00
	4	,817	2,860	,00	,00	,00	,00	,12	,07	,00	,13	,27	,03	,03
	5	,637	3,239	,00	,01	,00	,01	,41	,01	,00	,00	,44	,00	,00
	6	,515	3,604	,00	,00	,00	,02	,05	,04	,00	,00	,01	,25	,25
	7	,421	3,985	,00	,02	,00	,01	,03	,01	,00	,00	,04	,46	,46
	8	,272	4,954	,00	,12	,01	,00	,02	,00	,03	,01	,02	,15	,15
	9	,223	5,477	,00	,26	,00	,26	,18	,05	,01	,00	,03	,07	,07
	10	,135	7,035	,00	,05	,01	,26	,01	,61	,64	,55	,01	,00	,00
	11	,081	9,085	,02	,42	,39	,23	,11	,07	,04	,01	,03	,02	,02
	12	,020	18,128	,97	,11	,59	,20	,00	,06	,08	,05	,02	,00	,00

a. Dependent Variable: hodnoty; ekonomická levice praxe

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

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	5,137	,426		12,071	,000		
	vek	-,019	,005	-,184	-4,130	,000	,788	1,269
	nespokojenost s vnějšími podmínkami	-,029	,044	-,028	-,646	,519	,859	1,164
	prijem	,010	,005	,088	1,894	,059	,722	1,385
	učňovské vzdělání nebo sš bez maturity	,086	,209	,024	,411	,681	,471	2,124
	sš s maturitou	,185	,223	,049	,830	,407	,457	2,190
	vysokoškolské vzdělání	,277	,264	,056	1,048	,295	,555	1,802
	zajímá se o politiku	-,047	,149	-,013	-,314	,753	,878	1,139
	muž	,101	,145	,028	,693	,488	,927	1,079
	obec nad 5000 obyvatel	,057	,151	,015	,377	,706	,955	1,047
	chudi_nezam	-,779	,264	-,128	-2,952	,003	,828	1,207
	chudi	-,262	,198	-,058	-1,324	,186	,808	1,238
	bohati_nezam	-,110	,237	-,019	-,463	,644	,925	1,081

a. Dependent variable: hodnoty: ekonomická levice pravice

# Outlieři

- Např. pro případ č. 70 očekáváme, že nebude mít vyhraněný názor, ale přitom reálně jde o velmi levicového občana
- Podobně případ 105, ten je ale pravicový
- Můžeme vyřadit a zjistit, co to udělá s výsledky

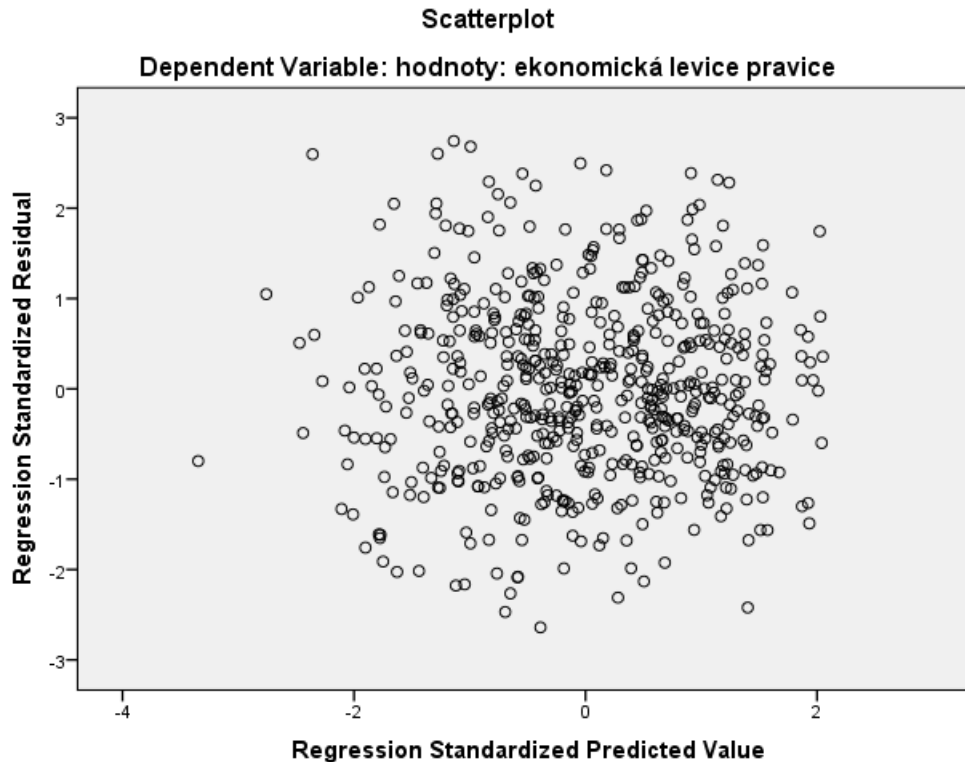
**Casewise Diagnostics<sup>a</sup>**

Case Number	Std. Residual	hodnoty: ekonomická levice pravice	Predicted Value	Residual
70	-2,177	,39	4,0863	-3,69304
105	2,283	9,23	5,3563	3,87237
127	-2,420	1,34	5,4433	-4,10519
156	2,015	60	3,0146	3,41904



# Homoskedascita

- V reziduích by neměl být žádný zřetelný vzorec



# heteroskedascita

- Příklad situace kdy homoskedascita není v pořádku

