Statistická analýza dat v psychologii I
Úvod do práce s jamovi
Seminář 5

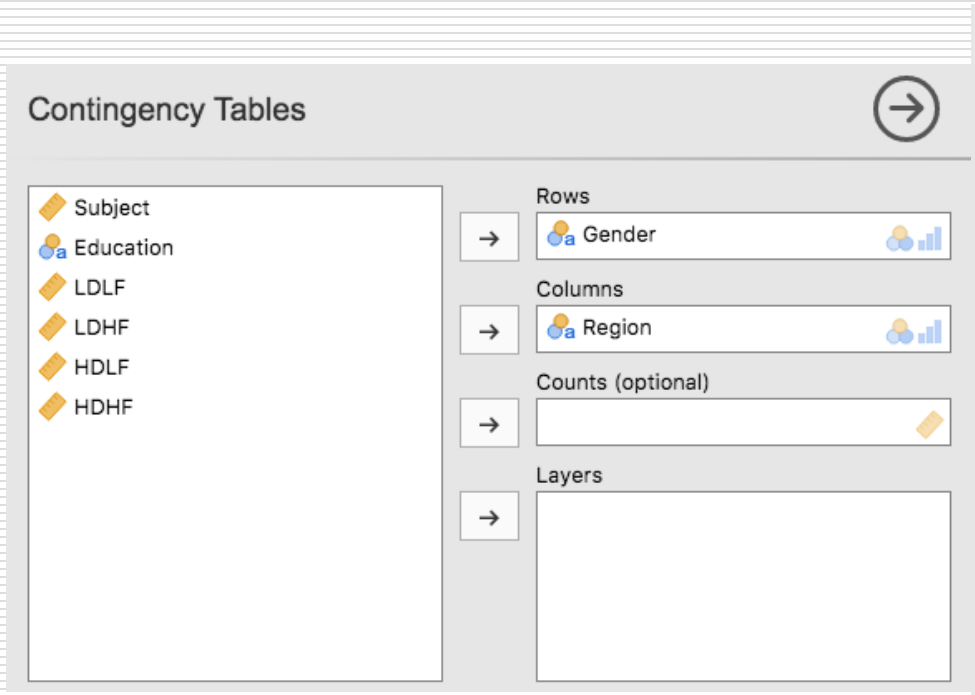
Vztahy mezi proměnnými, korelace, regrese

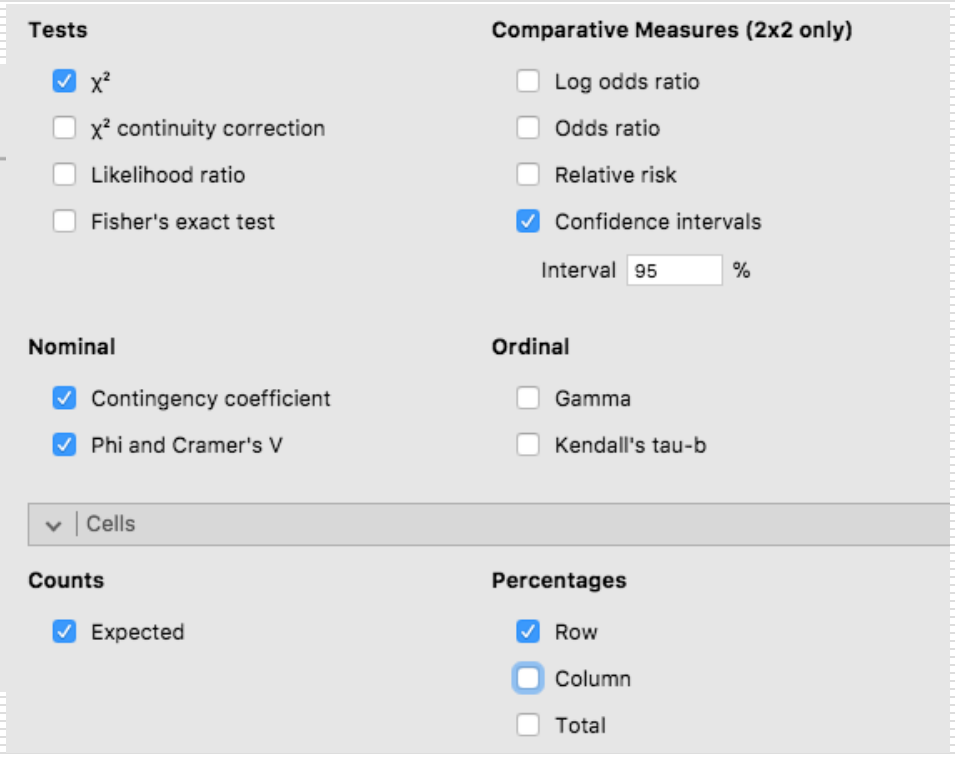
Obsah

- Kontingenční tabulka
- Korelace
  - Pearsonův, Spearmannův, Kendallův koeficient
  - Cronbachova alfa
- Lineární vztah
  - Bodový graf
- Lineární regrese

Kontingenční tabulka

- Frequencies – Independent Samples





Kontingenční tabulka

Contingency Tables

Gender		Region					Total
		Australia	Europe	North America	Other	South America	
Female	Observed	1	6	51	4	1	63
	Expected	0.685	6.16	52.7	2.74	0.685	
Male	Observed	0	3	26	0	0	29
	Expected	0.315	2.84	24.3	1.26	0.315	
Total	Observed	1	9	77	4	1	92
	Expected	1.000	9.00	77.0	4.00	1.000	

χ^2 Tests

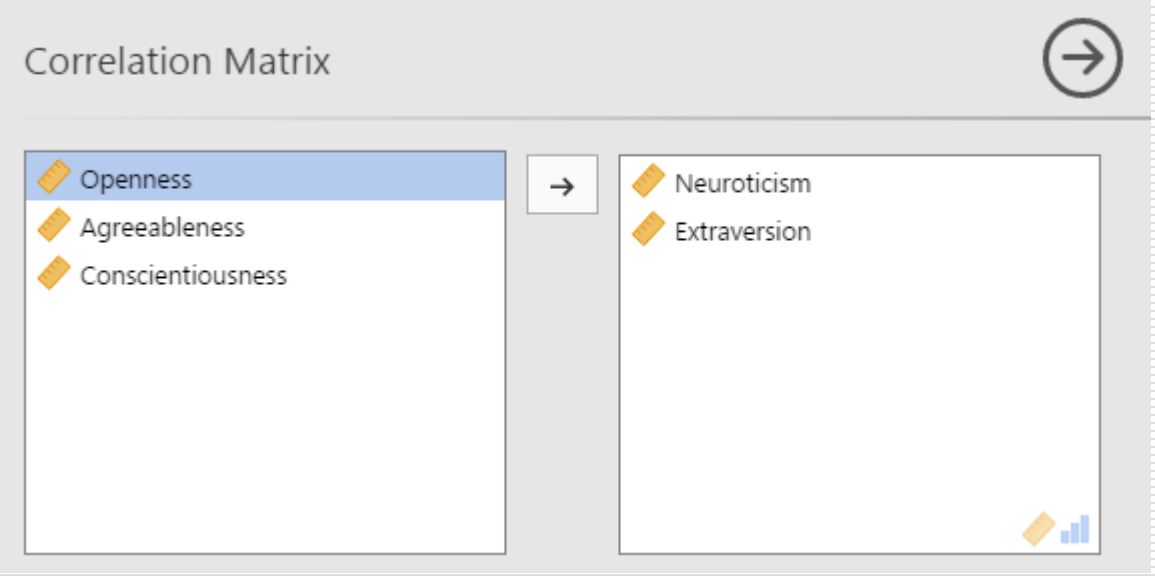
	Value	df	p
χ^2	2.96	4	0.565
N	92		

Nominal

	Value
Contingency coefficient	0.176
Phi-coefficient	NaN
Cramer's V	0.179

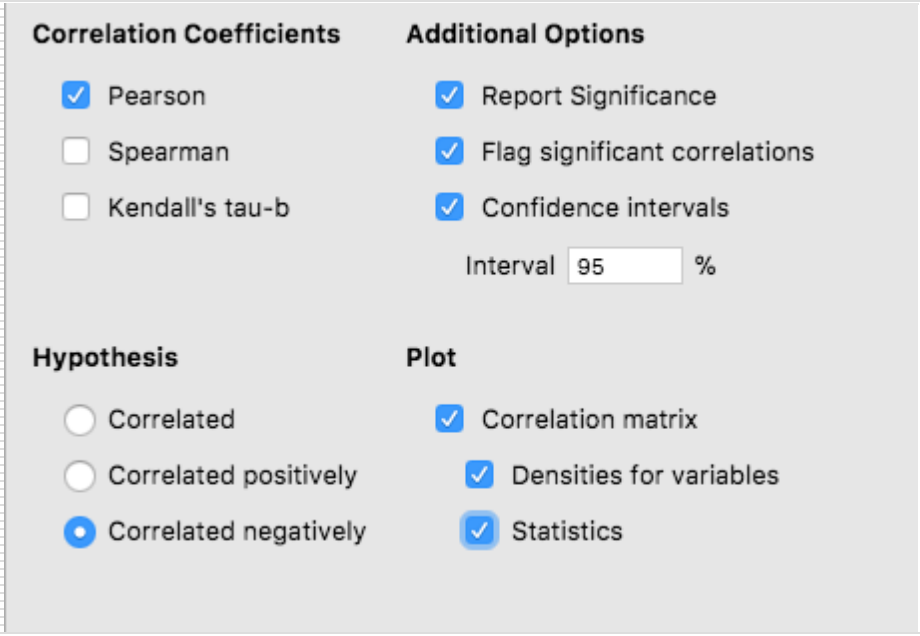
Korelace

- Regression – Correlation Matrix
  - Vybereme proměnné, jejichž souvislost nás zajímá



Korelace

- Kromě volby typu korelačního koeficientu je možné také testovat signifikanci korelací vč. uvedení intervalů spolehlivosti



Correlation Matrix

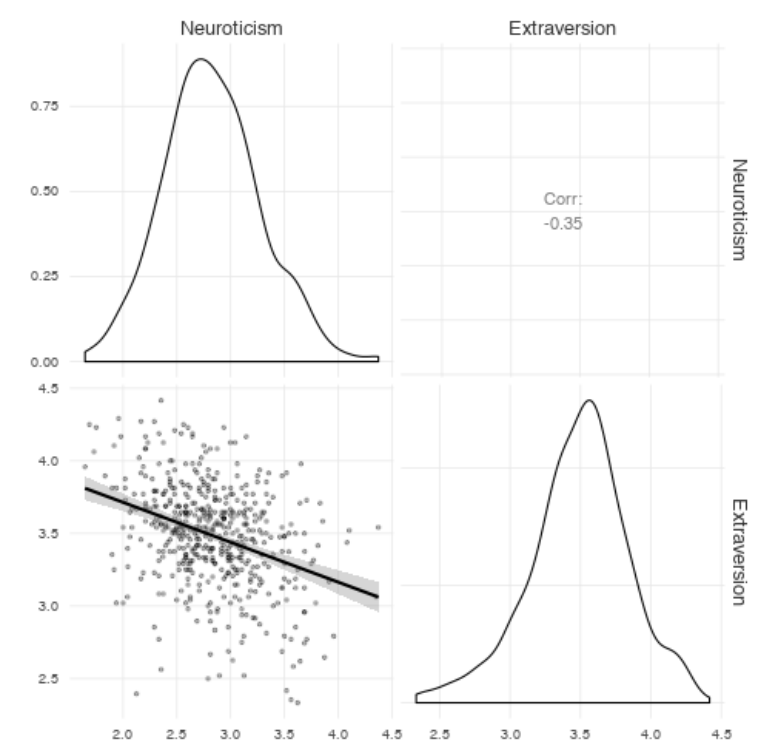
		Neuroticism	Extraversion
Neuroticism	Pearson's r	—	−0.350 ***
	p-value	—	<.001
	95% CI Upper	—	−0.284
	95% CI Lower	—	−1.000
Extraversion	Pearson's r		—
	p-value		—
	95% CI Upper		—
	95% CI Lower		—

Note. H_a is negative correlation

Note. * p < .05, ** p < .01, *** p < .001, one-tailed

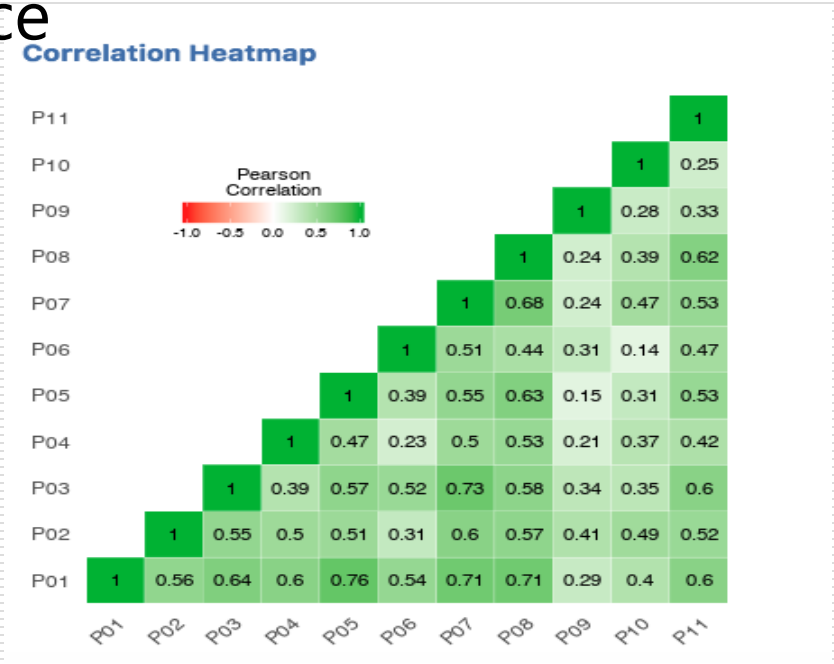
Korelace

- Kromě nulové hypotézy H: ρ = 0 lze v teorii odůvodněných případech testovat také hypotézy o "směru" korelace H: ρ > 0 a H: ρ < 0
- Dále je možné nechat vykreslit několik grafů



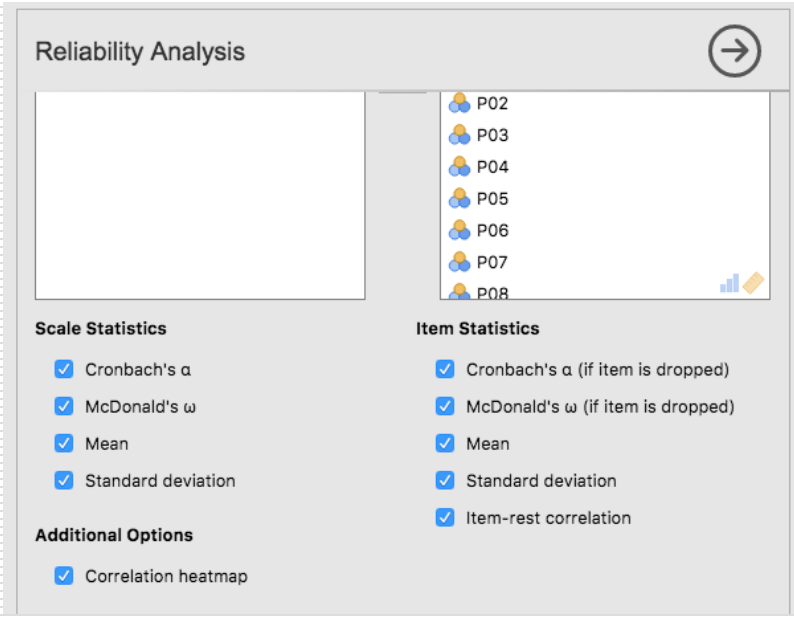
Korelace

- Factor – Reliability Analysis
  - *Correlation Heatmap* barevně vizualizuje vzájemné korelace



Cronbachova alfa

- Lze nechat zobrazit také Cronbachovu alfu a mnoho dalších ukazatelů, které patří spíše do Základů psychometriky



Reliability Analysis

Scale Reliability Statistics

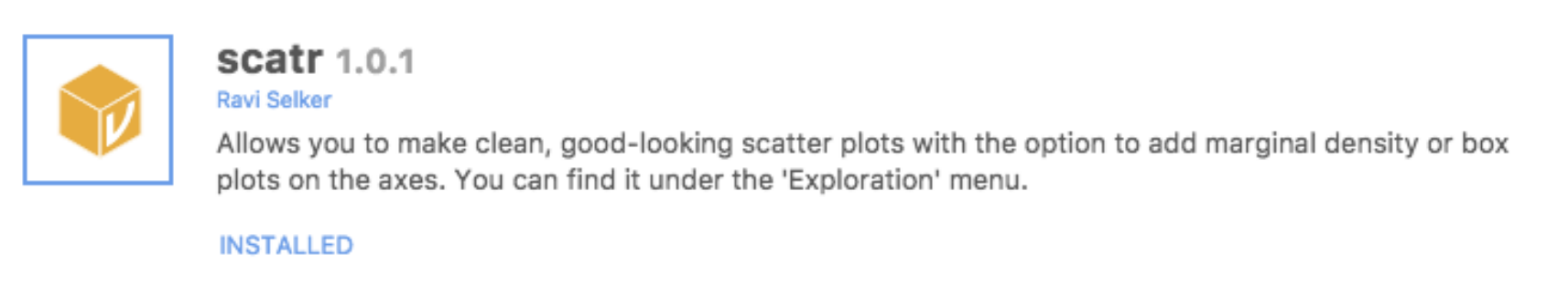
	mean	sd	Cronbach's α	McDonald's ω
scale	2.64	1.06	0.903	0.909

Item Reliability Statistics

				if item dropped	
	mean	sd	item-rest correlation	Cronbach's α	McDonald's ω
P01	1.88	1.62	0.835	0.883	0.891
P02	1.58	1.37	0.703	0.892	0.899
P03	3.81	1.32	0.735	0.890	0.896
P04	1.79	1.74	0.587	0.899	0.905
P05	2.42	1.82	0.685	0.893	0.899
P06	3.48	1.35	0.518	0.901	0.907
P07	3.28	1.56	0.782	0.887	0.893
P08	2.27	1.57	0.764	0.888	0.895
P09	3.36	1.28	0.365	0.908	0.914
P10	1.71	1.53	0.472	0.904	0.910
P11	3.48	1.07	0.679	0.895	0.900

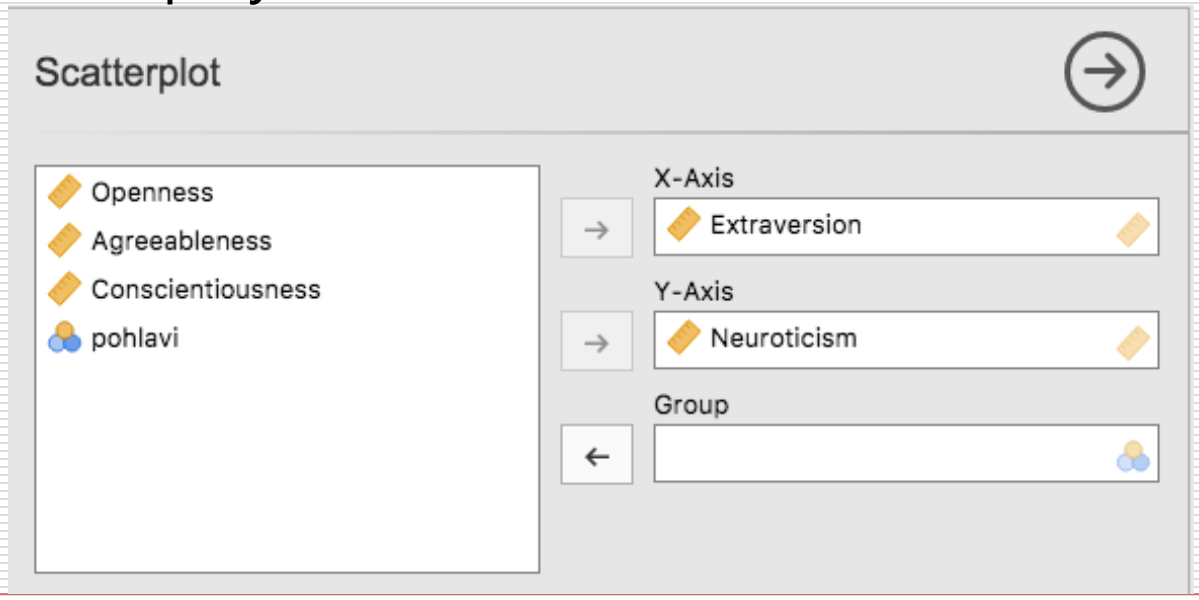
Lineární vztah – scatterplot

- Modules - jamovi library – scatr
  - V menu Descriptives poté naleznete možnost Scatterplot

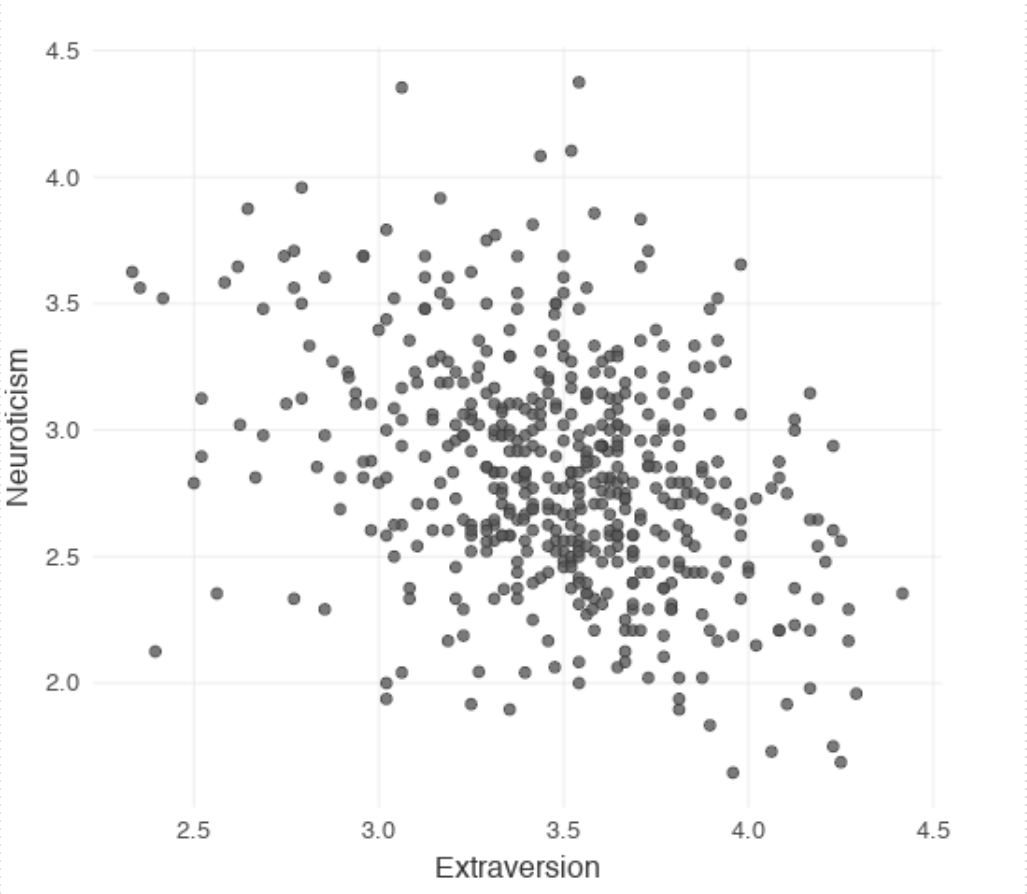


Scatterplot

- Graf lze vytvořit vybráním proměnné na X-ovou a Y-ovou osu
  - Okno Group umožňuje rozdělit vzorek na dvě skupiny

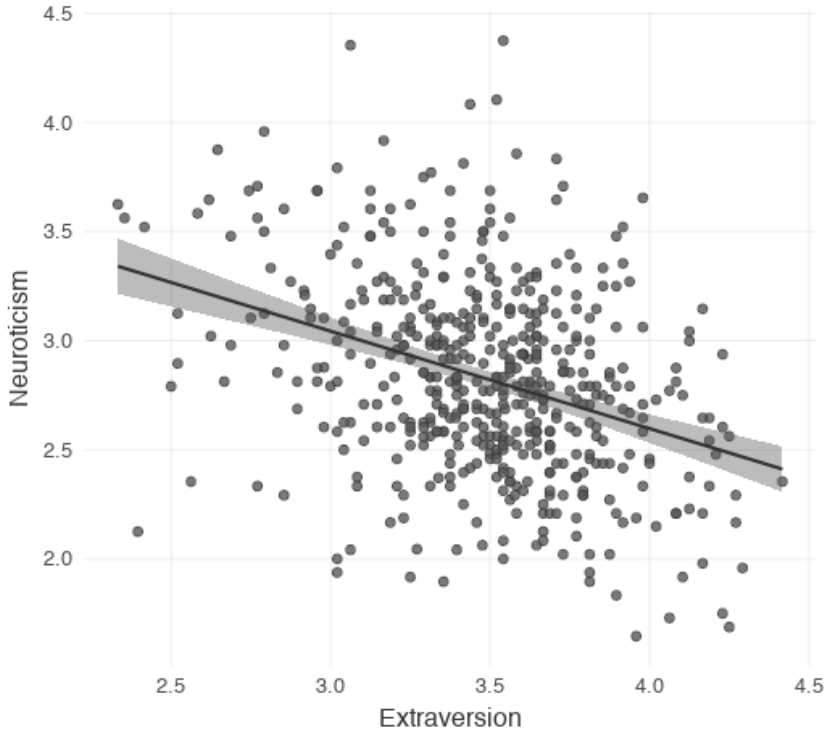


Scatterplot



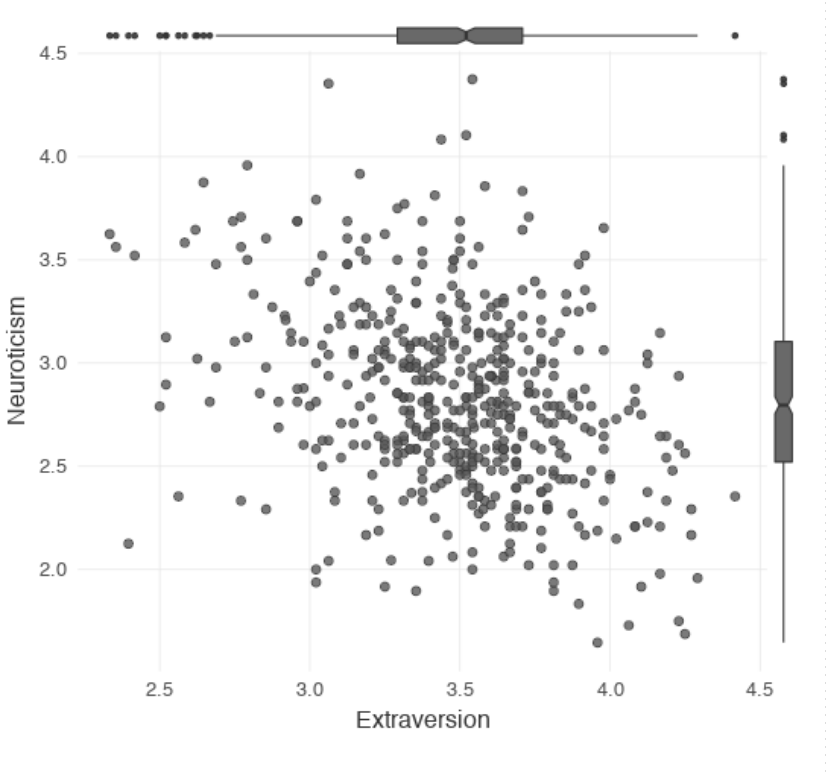
Scatterplot

- Dále lze do grafu přidat
  - Proložení regresní přímkou
  - Proložení regresní křivkou
    - Obě možnosti lze doplnit plochou symbolizující intervaly výběrové chyby



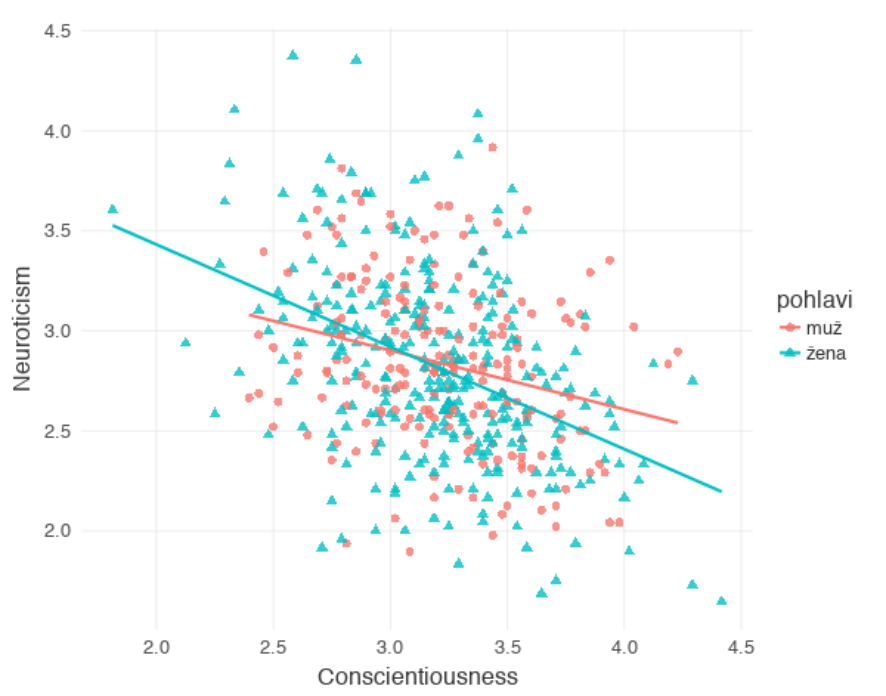
Scatterplot

- Na okraj grafu lze také doplnit
  - Graf hustoty dat
  - Boxploty



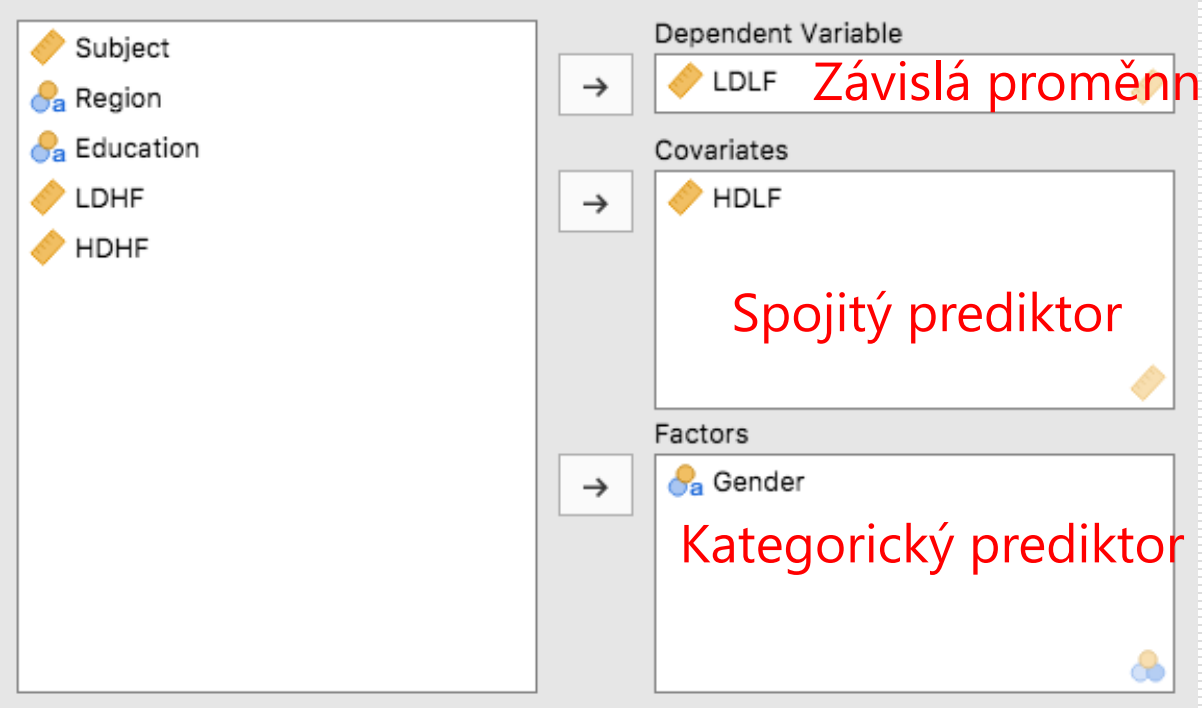
Scatterplot (navíc)

- PROTIP: Rozdělením proměnných na dvě skupiny lze graficky znázornit rozdílnost vztahu mezi proměnnými v závislosti na příslušnosti ke skupině (= moderace)



Lineární regrese

- Regression – Linear Regression
  - V Coefficients: Standardized estimate



Lineární regrese

$$Y = bX + a$$

Model Coefficients

Predictor	Estimate	SE	t	p	Stand. Estimate
Intercept	2.712	0.5936	4.57	<.001	
HDLF	0.442	0.0800	5.52	<.001	0.505

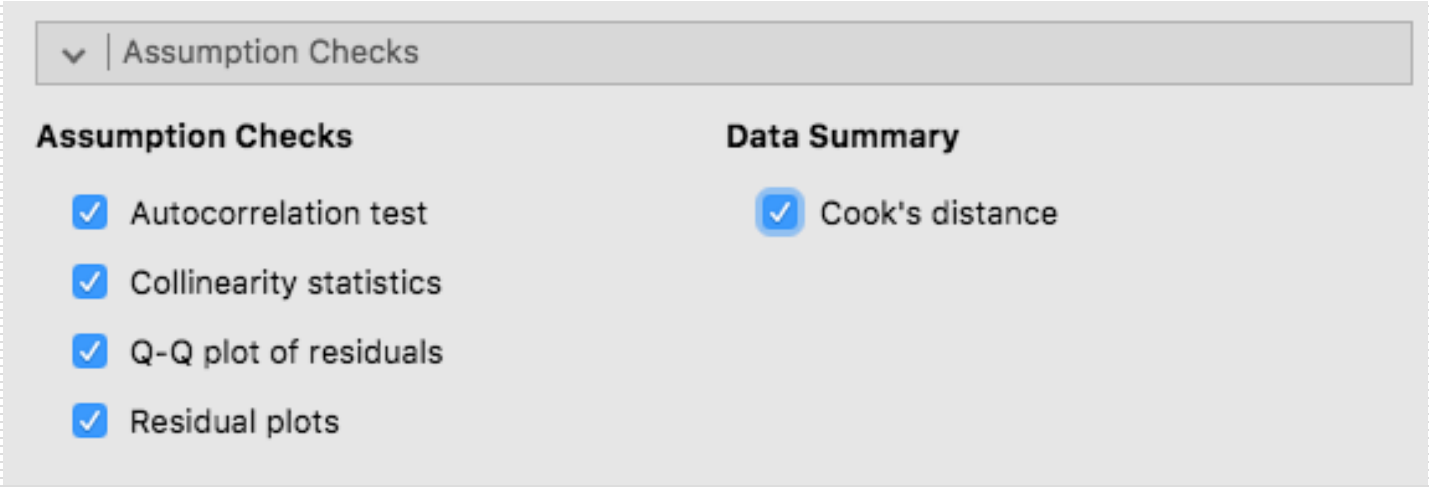
Koeficient determinance (R^2):

Model Fit Measures

Model	R	R^2
1	0.505	0.255

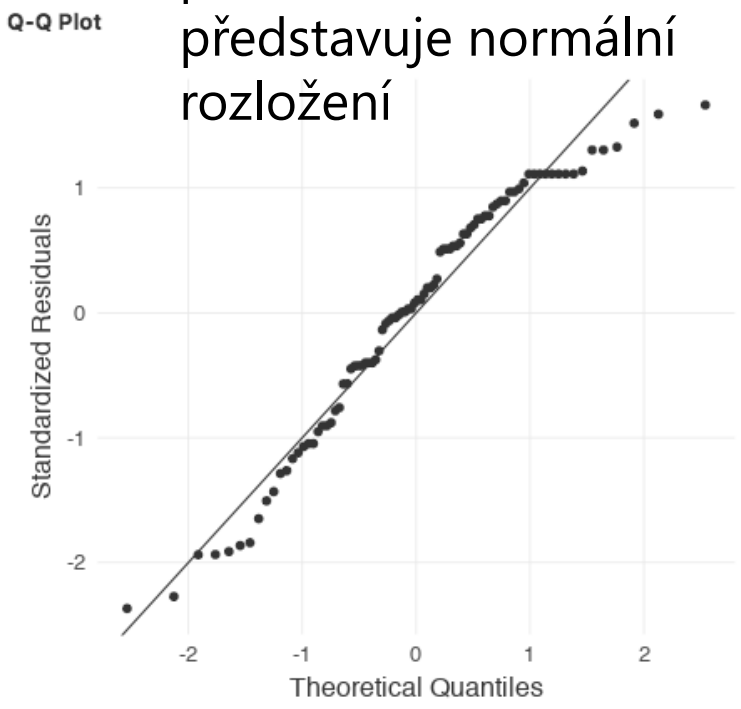
Lineární regrese (navíc)

- *Assumption Checks* (kontrola předpokladů regrese) nabízí mj. kontrolu normality residuí (*Q-Q plot of residuals*) a přítomnosti vlivných outlierů (*Cook's distance*)



Lineární regrese (navíc)

Chceme, aby residua byla co nejvíc na přímce, která představuje normální rozložení



Cook's Distance

			Range	
Mean	Median	SD	Min	Max
0.0111	0.00626	0.0134	5.91e−7	0.0612

Nechceme hodnoty větší než 3, vysoké cifry ukazují, že některý z případů je tzv. „vlivný“ = příliš ovlivňuje sklon regresní přímky

Lineární regrese (navíc)

- Jamovi skvěle ovládá také následující analýzy spojené s regresí, které potkáte ve Statistice II:
  - Moderaci a mediaci (modul *medmod*)
  - Logistickou regresi (*Regression – 2 Outcomes*)
  - Mixed modely (modul *Linear Models*)