

# CVIČENÍ Z ANALÝZY DAT SOCIOLOGICKÉHO VÝZKUMU

## 9. LEKCE

Jméno a příjmení:

Imatrik. ročník:

1.

- Jak velkou byste očekávali souvislost mezi důležitostí Boha (q33 v datech EVS ČR1999) a četností modliteb k Bohu mimo církevní obřady?
- Potvrdilo se vám toto očekávání? Jaký koeficient jste pro výpočet použili a proč?
- Pokud ano, o čem to svědčí?
- Co znamená, že má vypočtený korelační koeficient záporné znaménko?
- Je tato korelace signifikantní?

Řešení:

Analyze-Descriptive Statistics-Crosstabs-Statistics=Correlations

nebo

Analyze-Correlate-Bivariate-Spearman

Correlations

		Q33 Bůh - důležitost v životě	Q36 Modlení k Bohu mimo obřady
Spearman's rho	Q33 Bůh - důležitost v životě	Correlation Coefficient	1,000
		Sig. (2-tailed)	,000
		N	1880
	Q36 Modlení k Bohu mimo obřady	Correlation Coefficient	-,784*
		Sig. (2-tailed)	,000
		N	1857

\*\* . Correlation is significant at the .01 level (2-tailed).

- Očekávali jsme vysokou korelaci a b) ona se potvrdila. Spearmana, jedna z proměnných je ordinální povahy. c) Svědčí to o tom, že respondenti jsou ve svých odpovědích consistentní. d) Záporné znaménko znamená v daném případě kladnou korelaci: čím je Bůh důležitější, tím častěji se respondent modlí k Bohu i mimo církevní obřady. e) Ano, korelace je signifikantní.

2.

- Do jaké míry souvisí postoje v baterii q54a – q54h se vzděláním respondenta (ISCED1)? Výpočet udělejte jako obdélníkovou matici, v níž vzdělání je nezávisle proměnnou. Formát výstupního výpočtu přetvořte tak, aby vzdělání bylo umístěno ve sloupci a závisle proměnné v řádcích.
- Jaký koeficient jste pro výpočet použili?
- S jakou proměnnou v dané baterii vzdělání nejvíce koreluje? Jaký je příslušný koeficient determinace?
- A s jakou proměnnou naopak vzdělání koreluje nejméně?
- Proč jsou některé korelace kladné a některé záporné?
- Proč jsou korelace tak rozdílné?
- Vyjdou výsledky stejně, pokud proměnnou vzdělání měřenou jako znak *ISCED1* (se šesti kategoriemi) nahradíte proměnnou *vzdelání*, která má kategorie pouze čtyři?)

Řešení:

Omlouvám se, k tomuto úkolu jste ve svém souboru neměli proměnnou ISCED1, takže stačí pracovat s proměnnou vzdělání

Analyze-Descriptive Statistics-Crosstabs-Statistics=Correlations

nebo

Analyze-Correlate-Bivariate-Spearman

Pokud ale chceme, aby vzdělání bylo ve sloupci, musíme použít syntaxe:

CORRELATIONS

/VARIABLES=q54a q54b q54c q54d q54e q54f q54g q54h with ISCED1

/PRINT=TWOTAIL NOSIG

/MISSING=PAIRWISE.

Correlations

	Spearman's rho		
	ISCED1		
	Correlation Coefficient	Sig. (2-tailed)	N
Q54A Odpovědnost za jednotlivce má:	-,192**	,000	1860
Q54B Nezaměstnaní nabízené zaměstnání:	,024	,312	1850
Q54C Konkurence je:	-,150**	,000	1848
Q54D Stát by měl:	-,170**	,000	1834
Q54E Příjmy by měly být:	,205**	,000	1851
Q54F Rozšiřovat/ zachovat by se mělo obchod. a prům.:	-,187**	,000	1823
Q54G Za důchodové zabezp. má být zodpovědný:	-,173**	,000	1851
Q54H Za bydlení má být zodpovědný:	-,135**	,000	1860

\*\* - Correlation is significant at the .01 level (2-tailed).

a) Souvisí do určité míry se všemi, ale korelační koeficienty jsou poměrně nízké a v jednom případě i nesignifikanční, b) Použili jsme Spearmana, protože se jedná o ordinální proměnné, z nichž jedna je dlouhá, c) Nejvíce koreluje s E, tj. Příjmy by měly být vyrovnanější nebo naopak víc motivační, koeficient determinace je  $0,205^2$ , tj. 0,042. Vzdělání tedy určuje jen 4 % variability této proměnné. d) Nejméně koreluje s B, tj. Nezaměstnaní nabízené zaměstnání mají povinnost přijmout nebo ho mohou odmítnout e) Záleží na tom, zda se jedná o přímou nebo nepřímou úměru, f) Jsou různě postavené otázky a otočené stupnice, g) Korelace vycházejí podobně.

Correlations

	Spearman's rho		
	VZDELÁNÍ kategorizace q94		
	Correlation Coefficient	Sig. (2-tailed)	N
Q54A Odpovědnost za jednotlivce má:	-,172**	,000	1864
Q54B Nezaměstnaní nabízené zaměstnání:	,048*	,039	1855
Q54C Konkurence je:	-,132**	,000	1850
Q54D Stát by měl:	-,181**	,000	1838
Q54E Příjmy by měly být:	,219**	,000	1852
Q54F Rozšiřovat/ zachovat by se mělo obchod. a prům.:	-,171**	,000	1830
Q54G Za důchodové zabezp. má být zodpovědný:	-,162**	,000	1853
Q54H Za bydlení má být zodpovědný:	-,117**	,000	1865

\*\* - Correlation is significant at the .01 level (2-tailed).

\* - Correlation is significant at the .05 level (2-tailed).

3.

- Testujte nulovou hypotézu, že mezi hrdostí na to být občanem ČR (q67) a preferencí místa, k němuž respondenti pocítují největší příslušnost není žádná souvislost (q71). Jaký koeficient zvolíte a proč?
- Jak danou souvislost vysvětlíte?
- Kolik bylo v souboru „evropo“ a „světoobčanů“, to je lidí pocítujících největší příslušnost k Evropě a ke světu? Kolik je to procent z celého souboru?

Řešení:

Analyze-Descriptive statistics-Crosstabs-Correlations-Statistics=Correlations nebo Kendallovo tau-c, Cells-Percentage

Q67 Nejsilnější pocit příslušnosti \* Q71 Hrdost na občanství ČR Crosstabulation

			Q71 Hrdost na občanství ČR				Total
			1 velmi hrdý/á	2 dosít hrdý/á	3 ne moc hrdý/á	4 vůbec ne hrdý/á	
Q67 Nejsilnější pocit příslušnosti	1 místo, obec, kde žije	Count	205	430	136	13	784
		Column %	42,6%	42,8%	43,7%	37,1%	42,8%
	2 Čechy, Morava, Slezsko	Count	63	131	50	7	251
		Column %	13,1%	13,0%	16,1%	20,0%	13,7%
	3 ČR	Count	190	382	76	4	652
	Column %	39,5%	38,0%	24,4%	11,4%	35,6%	
	4 Evropa	Count	9	16	18	4	47
		Column %	1,9%	1,6%	5,8%	11,4%	2,6%
	5 svět jako celek	Count	14	46	31	7	98
		Column %	2,9%	4,6%	10,0%	20,0%	5,3%
Total		Count	481	1005	311	35	1832
		Column %	100,0%	100,0%	100,0%	100,0%	100,0%

Symmetric Measures

		Value	Asymp. Std. Error <sup>a</sup>	Approx. T <sup>b</sup>	Approx. Sig.
Nominal by Nominal	Phi	,214			,000
	Cramer's V	,124			,000
Ordinal by Ordinal	Kendall's tau-c	,015	,018	,829	,407
	Spearman Correlation	,020	,024	,860	,390 <sup>c</sup>
	Pearson's R	,052	,025	2,208	,027 <sup>c</sup>
N of Valid Cases		1832			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

c. Based on normal approximation.

a) Je možné zvolit např. Kendala nebo Spearmana, protože tyto proměnné lze seřadit. Korelace je ale nízká, navíc nesignifikantní. Nízkou korelaci dokládá i rozložení dat v tabulce, viz první řádek dat. Ať je člověk hrdý více nebo méně, pocítuje nejsilnější příslušnost k obci, v níž žije.

- b) Nejvíce lidí pocítuje nejsilnější příslušnost k obci, hrdost na občanství je řádově odlišná.  
 c) K Evropě 47, ke světu 98, celkem tedy  $145/1832 = 8 \%$ .

**4.**

Sledujte, jak se mění velikost korelačních koeficientů a jejich signifikance v závislosti na velikosti souboru. Zjistěte nejdříve, jaká je korelace (měřená Pearsonem) mezi délkou rozhovoru (proměnná *c51*) a věkem (proměnná *vek*) v celém souboru. poté udělejte náhodný výběr 600 respondentů z 1900 a spočítejte tutéž korelaci. Pak udělejte náhodný výběr 200 respondentů, dále 65 respondentů a 30 respondentů. V každém náhodném výběru spočítejte korelaci a zjistěte signifikanci. Výsledky запиšte do tabulky. Budou výsledky u všech studentů stejné, nebo se budou lišit. pokud ano, tak proč?

**Data-Select cases-Random sample of cases**

**Analyze-Correlate-Bivariate-Pearson**

**nebo**

**Analyze-Descriptive Statistics-Crosstabs-Statistics=Correlations**

**Tab.:** Korelace mezi délkou rozhovoru a věkem respondenta + signifikance

N = celý soubor		N = 600		N = 200		N = 65		N = 30	
r	sign.	r	sign.	r	sign.	r	sign.	r	sign.
0,15	0,000	0,27	0,000	0,20	0,006	0,13	0,286	0,38	0,032

Pozor: Výsledky se budou u jednotlivých studentů lišit, neboť náhodné výběry budou u každého studenta odlišné!

Korelace se s velikostí souboru snižuje.