



Means – průměry

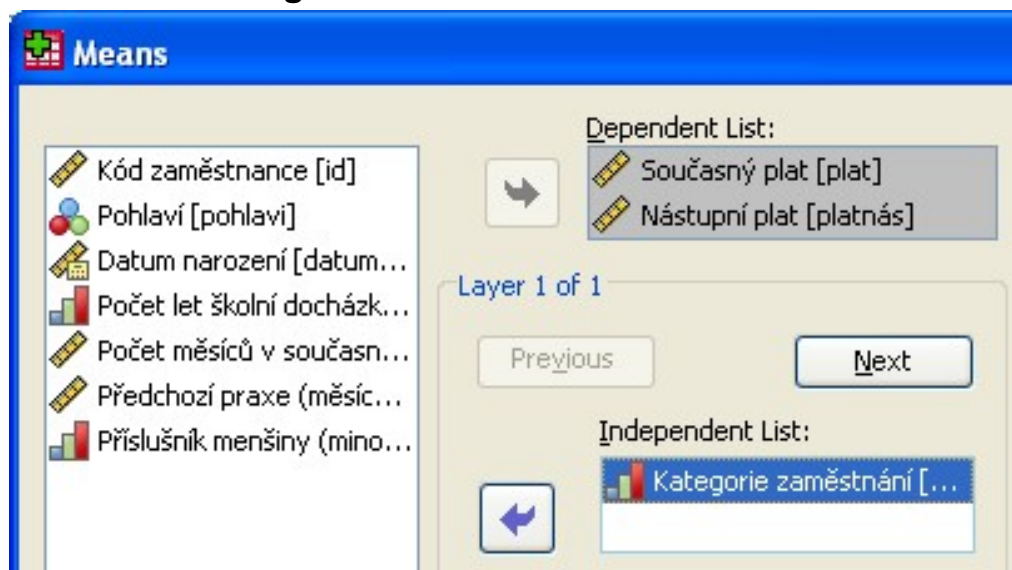
Tabulky průměrů

Procedura *Means* se užívá pro porovnání průměrů a dalších popisných statistik v podskupinách. Analyzovat můžeme současně několik číselných proměnných. Zadáme-li více kategorizovaných faktorů určujících podskupiny, lze výstup organizovat buď v samostatných tabulkách pro každou třídící proměnnou zvlášť, nebo jako třídění vyššího stupně. K dispozici je rovněž tabulka analýzy rozptylu (ANOVA) a test lineární závislosti průměrů pro první stupeň třídění.

Volání procedury v IBM SPSS Statistics

Analyze → Compare Means → Means

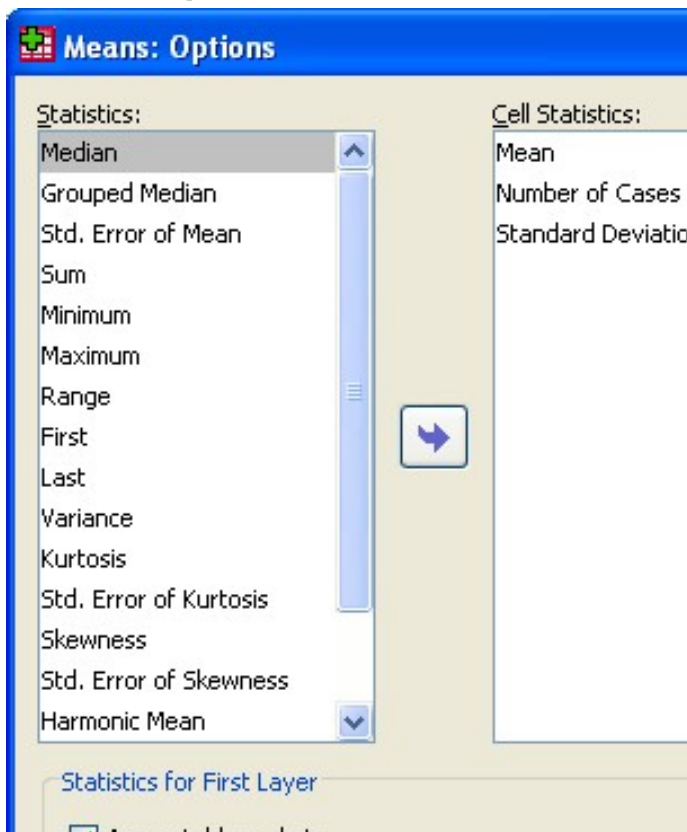
Nastavení dialogu



- Do pole *Dependent list* přeneseme číselné proměnné, jejichž popisné statistiky počítáme.
- Proměnné určující skupiny zadáme do pole *Independent list*. Jestliže sem vložíme více proměnných současně, vytvoří se pro každou z nich samostatná tabulka.
- Pokud potřebujeme získat tabulku vyššího stupně třídění, využijeme tlačítka *Previous* a *Next*, která přepínají mezi jednotlivými vrstvami třídění.

Chceme-li například vytvořit tabulku třídění druhého stupně podle *kategorí zaměstnání* a *pohlaví*, přeneseme do pole *Independent List* proměnnou *kategorie zaměstnání*, stiskneme tlačítko *Next* a ve druhé vrstvě zadáme opět do pole *Independent List* proměnnou *pohlaví*.

Tlačítko *Options*



Pomocí tlačítka *Options* volíme statistiky, které se zobrazí ve výstupové tabulce, a případně doplňující statistické testy.

Statistics, Cell Statistics (Zadávání statistik)

V levé části okna (*Statistics*) se nachází široký výběr popisných statistik. Požadované statistiky přeneseme šipkou do pole *Cell Statistics*.

Anova table and eta (Analýza rozptylu)

Po zaškrtnutí políčka *Anova table and eta* se na výstupu kromě tabulky popisných statistik zobrazí také tabulka analýzy rozptylu (ANOVA) pro třídící proměnné uvedené v první vrstvě. Tabulka umožňuje posoudit na základě statistického testu, zda jsou rozdíly průměrů ve skupinách významné. Koeficient *eta*, který nabývá hodnot od nuly do jedné, pak vyjadřuje míru závislosti třídících a analyzovaných proměnných (čím větší *eta*, tím silnější závislost).

Test for linearity (Test linearity)

Při označení políčka *Test for linearity* se modifikují tabulky analýzy rozptylu, kde je navíc testována lineární závislost průměrů na skupinách první vrstvy. Kromě toho se spočítá také Pearsonův korelační koeficient. Tyto analýzy mají smysl v případě, že třídící proměnná je ordinální (kategorie jsou uspořádané).

Výstupy

Přehled o počtu platných a chybějících případů

Case Processing Summary						
	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Současný plat * Kategorie zaměstnání	474	100.0%	0	.0%	474	100.0%
Nástupní plat * Kategorie zaměstnání	474	100.0%	0	.0%	474	100.0%

Informace o počtu platných a chybějících hodnot pro uvedené kombinace proměnných.

Přehled statistik

Report			
Kategorie zaměstnání		Současný plat	Nástupní plat
úředník	Mean	\$27,838.54	\$14,096.05
	N	363	363
	Std. Deviation	\$7,567.995	\$2,907.474
manuál. zaměstnání	Mean	\$30,938.89	\$15,077.78
	N	27	27
	Std. Deviation	\$2,114.616	\$1,341.235
manažer	Mean	\$63,977.80	\$30,257.86
	N	84	84
	Std. Deviation	\$18,244.776	\$9,980.979
Total	Mean	\$34,419.57	\$17,016.09
	N	474	474
	Std. Deviation	\$17,075.661	\$7,870.638

Tabulka zobrazuje popisné statistiky zadaných proměnných uvnitř definovaných skupin. Strukturu tabulky můžeme dále upravit pivotováním. Pomocí této techniky lze modifikovat hierarchii sloupců, řádků a vrstev tak, aby tabulka přehledně zobrazovala požadovanou informaci.

Analýza rozptylu

ANOVA Table						
		Sum of Squares	df	Mean Square	F	Sig.
Současný plat * Kategorie zaměstnání	Between Groups (Combined)	89438483926	2	44719241963	434.5	.000
	Within Groups	48478011510	471	102925714		
	Total	137916495436	473			
Nástupní plat * Kategorie zaměstnání	Between Groups (Combined)	17925544532	2	8962772266	371.1	.000
	Within Groups	11375360433	471	24151508		
	Total	29300904965	473			

Listy procedur IBM SPSS Statistics

Tabulku analýzy rozptylu využijeme pro testování, zda kategorie zaměstnání je významným faktorem ovlivňujícím současný nebo nástupní plat zaměstnance. Nulová hypotéza je formulována tak, že průměrný plat na celé populaci je u všech kategorií zaměstnání stejný. Vzhledem k tomu, že hodnota signifikancí v posledním sloupci tabulky je v obou případech velmi malá, můžeme na 95% hladině spolehlivosti zamítnout nulovou hypotézu.

Measures of Association

	Eta	Eta Squared
Současný plat * Kategorie zaměstnání	.805	.648
Nástupní plat * Kategorie zaměstnání	.782	.612

Ze sloupce *Eta Squared* zjistíme, že *kategorie zaměstnání* vysvětluje variabilitu proměnné *současný plat* z 64,8 %, *nástupní plat* z 61,2 %.

Testování linearity

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Současný plat * Kategorie zaměstnání	Between Groups	(Combined)	89438483926	2	44719241963	434.5	.000
		Linearity	83933110461	1	83933110461	815.5	.000
		Deviation from Linearity	5505373465	1	5505373465	53.5	.000
	Within Groups		48478011510	471	102925714		
	Total		137916495436	473			
Nástupní plat * Kategorie zaměstnání	Between Groups	(Combined)	17925544532	2	8962772266	371.1	.000
		Linearity	16687316491	1	16687316491	690.9	.000
		Deviation from Linearity	1238228041	1	1238228041	51.3	.000
	Within Groups		11375360433	471	24151508		
	Total		29300904965	473			

Budeme-li považovat kategorie zaměstnání za ordinální proměnnou, můžeme rovněž využít test linearity pro posouzení, zda v datech existuje lineární vztah mezi současným nebo nástupním platem a kategoriemi zaměstnání. V našem případě jsou všechny hodnoty signifikancí v posledním sloupci velice nízké, což napovídá, že lineární závislost v datech existuje, je zde však také výrazná odchylka od linearity.

Measures of Association

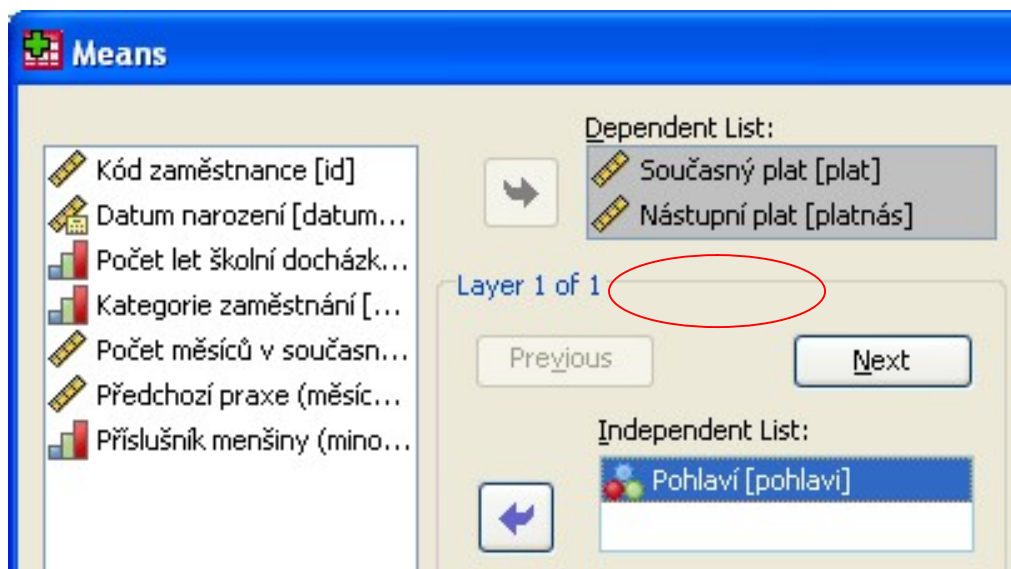
	R	R Squared	Eta	Eta Squared
Současný plat * Kategorie zaměstnání	.780	.609	.805	.648
Nástupní plat * Kategorie zaměstnání	.755	.570	.782	.612

Ve sloupci tabulky označeném *R* jsou uvedeny hodnoty Pearsonova lineárního korelačního koeficientu daných proměnných. Následující sloupec *R Squared* udává druhou mocninu tohoto koeficientu. Jedná se o míru, která charakterizuje kvalitu lineárního modelu. (Přesněji *R Squared*, někdy též *koeficient determinace*, udává procento variability vysvětlené pomocí lineárního regresního vztahu.)

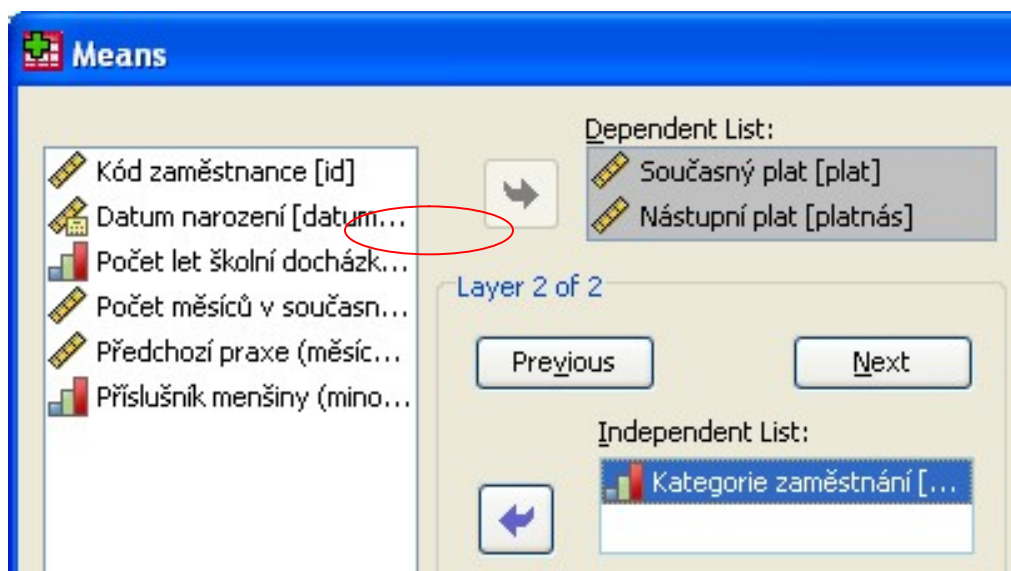
Ukázka zadávání tabulky

Potřebujeme vytvořit tabulku obsahující průměry a mediány proměnných *Současný plat* a *Nástupní plat* tříděnou nejprve podle *Pohlaví* respondenta a dále podle *Kategorie zaměstnání*.

V okně procedury *Means* zadáme do pole *Dependent List* proměnné *Současný plat* a *Nástupní plat*, do pole *Independent List* proměnnou *Pohlaví*.

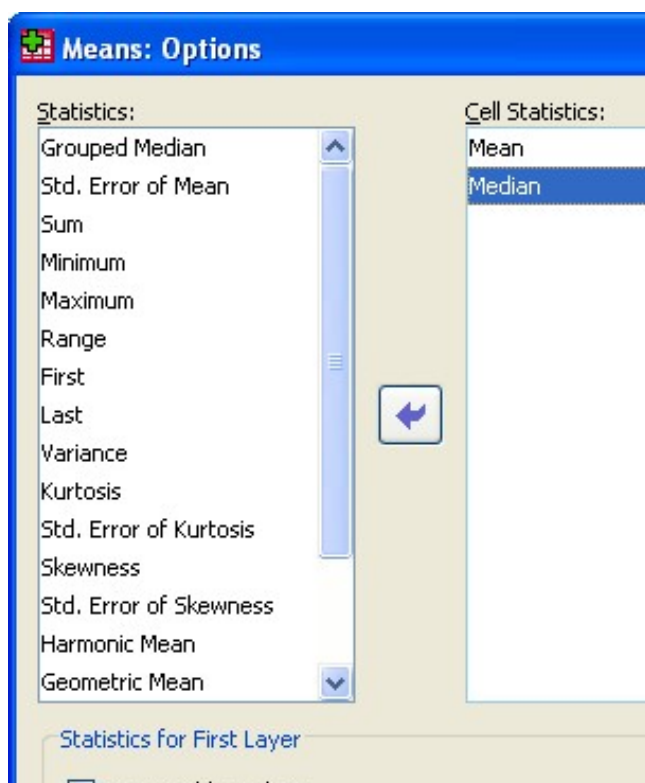


Dále stiskneme tlačítko *Next* a do pole *Independent List* přeneseme proměnnou *Kategorie zaměstnání*.



Stiskneme tlačítko *Options* a do pole *Cell Statistics* přeneseme požadované statistiky.

Listy procedur IBM SPSS Statistics



Po stisknutí tlačítek *Continue* a *OK* dostaneme následující tabulku.

Report

Pohlaví	Kategorie zaměstnání		Současný plat	Nástupní plat
žena	úředník	Mean	\$25,003.69	\$12,750.75
		Median	\$24,000.00	\$12,000.00
	manažer	Mean	\$47,213.50	\$20,121.00
		Median	\$45,187.50	\$19,125.00
	Total	Mean	\$26,031.92	\$13,091.97
		Median	\$24,300.00	\$12,375.00
muž	úředník	Mean	\$31,558.15	\$15,861.21
		Median	\$29,850.00	\$15,750.00
	manuál. zaměstnání	Mean	\$30,938.89	\$15,077.78
		Median	\$30,750.00	\$15,750.00
	manažer	Mean	\$66,243.24	\$31,627.70
		Median	\$63,750.00	\$30,000.00
	Total	Mean	\$41,441.78	\$20,301.40
		Median	\$32,850.00	\$15,750.00
Total	úředník	Mean	\$27,838.54	\$14,096.05
		Median	\$26,550.00	\$14,100.00
	manuál. zaměstnání	Mean	\$30,938.89	\$15,077.78
		Median	\$30,750.00	\$15,750.00
	manažer	Mean	\$63,977.80	\$30,257.86
		Median	\$60,500.00	\$28,740.00
	Total	Mean	\$34,419.57	\$17,016.09
		Median	\$28,875.00	\$15,000.00