



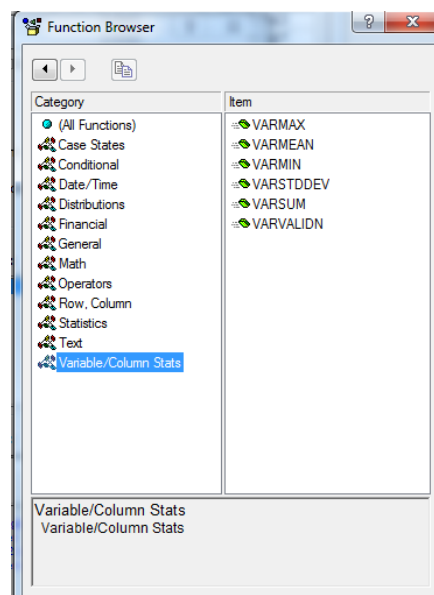
StatSoft

Statistiky sloupců jako funkce

Rádi bychom představili několik nových funkcí, které výrazně rozšiřují možnosti programu *STATISTICA*. Troufáme si říct, že sami budete rádi, že je nyní máte.

Kde je najdeme

Pojďme hned na to a ukažme si, kde nové funkce nalézt. Jedná se o funkce pro proměnné přístupné v tabulce *STATISTICA*. Otevřeme dialog nějaké například nové proměnné a klepneme na tlačítko *Funkcemi* (v anglické verzi *Functions*). Ve výčtu funkcí je speciální kategorie *Statistiky proměnné/sloupce* (v anglické verzi nazvaná *Variable/Column Stats*). Vidíme zde několik funkcí, které začínají na „VAR“.



Co umí

Těchto několik funkcí je určeno pro výpočet statistik pro vybraný sloupec/proměnnou. Pokud tedy napíšeme do prostoru pro dlouhý název vzorec „=VARMEAN(v1)“, pak bude tato proměnná vyplněna průměrem hodnot v prvním sloupci. Kromě průměru máte k dispozici součet, počet platných pozorování, maximum, minimum a směrodatnou odchylku.

Rozdíl mezi klasickými statistikami a statistikami pro proměnné

	1 první proměnná	2 funkce =mean(v1)	3 funkce =varmean(v1)
1	1	1	5,5
2	2	2	5,5
3	4	4	5,5
4	8	8	5,5
5	9	9	5,5
6	10	10	5,5
7	2	2	5,5
8	3	3	5,5
9	8	8	5,5
10	8	8	5,5

Ukažme si rozdíl na funkcích *mean* a *varmean*. Obě znamenají průměr, nicméně každá trochu něčeho jiného. Pokud zadáme pro funkce statistik stejné argumenty například první proměnnou, tedy „=mean(v1)“ a „=varmean(v1)“, získáme tabulku vlevo. Vidíme, že výsledky jsou rozdílné. Zatímco ve druhé proměnné se provedl průměr jedné hodnoty z prvního sloupce ve stejném řádku - jelikož argumentem funkce *mean* je jen první proměnná, ve třetí proměnné je všude průměr všech deseti hodnot první proměnné. Využití funkce *mean* je ale v reálu úplně jiné, protože dělá průměr po řádcích. Podívejme se na rozumnější

využití v tabulce vpravo. Ve třetí proměnné je průměr prvních dvou pro každý řádek zvlášť. Třetí proměnná tedy obsahuje průměry a ne jen jeden průměr jako tomu bylo u funkce „=varmean(v1)“.

	1 první proměnná	2 druhá proměnná	3 funkce =mean(v1:v2)
1	1	3	2
2	2	2	2
3	4	10	7
4	8	5	6,5
5	9	7	8
6	10	4	7
7	2	7	4,5
8	3	8	5,5
9	8	5	6,5
10	8	1	4,5

Oba principy jsou jistě využitelné, nicméně každý pro trochu jinou situaci. Možnost mít k dispozici obě varianty je jistě výhodná.

Použití statistik sloupců

Statistiky sloupců můžete využít nejrůznějšími způsoby. Nejjednodušším a běžným použitím je například potřeba odečíst od dané proměnné charakteristiku (statistiku) nějaké proměnné. Typickým příkladem může být například standardizace sloupce (může být sice provedena jinak a jednodušeji, ale pomocí funkcí výše nadefinovaných to lze také). Pokud bychom chtěli standardizovanou veličinu $v1$, uvedli bychom vzorec: „ $=(v1-VARMEAN(v1))/VARSTDDEV(v1)$ “. Dalším příkladem může být náhrada velkých hodnot maximem, například takto: „ $=if(v1>124;VARMAX(v1);v1)$ “. Jistě se i u Vás najde mnoho dalších případů, kdy potřebujete využít statistiku nějaké proměnné. Důležité je vědět, že tyto funkce v softwaru jsou a že je můžete využívat.