

StatSoft Vzorce

Jistě se Vám již stalo, že data, která máte přímo k dispozici, sama o sobě nestačí – potřebujete je nějak upravit, vypočítat z nich nějaké další proměnné, provést nějaké transformace, ... Jinak řečeno, neobejdete se bez vzorců.

V tomto krátkém článku si ukážeme základní možnosti, jak využívat množství funkcí, které *STATISTICA* poskytuje, k tomu si také ukážeme s tím související triky a tipy.

Vzorce

Podobně jako například v MS Excel, i v programu *STATISTICA* máme možnost dopočítávat hodnoty v datové tabulce pomocí nejrůznějších funkcí.

Vzorce jsou z důvodu statistického využití dat koncipovány jako vzorce pro sloupec (proměnnou) – hodnoty v jednom sloupci jsou tedy vypočteny stejným vzorcem, není tedy možné zadat například vzorec pouze pro jednu konkrétní buňku. Z logiky statistických dat: každý řádek souboru jsou nějaká pozorování, měření či hodnoty pro danou jednotku, všechny jednotky mají stejný typ měření a nemá tedy smysl uvažovat vzorce jen pro část z nich. Ve vzorcích můžete využívat nejrůznější funkce, které si můžete vybrat ze seznamu funkcí. Můžete využít funkce matematické, logické operátory, podmínky, [funkce pro práci s datumem](#), stejně jako generátory náhodných čísel či funkce pro [výpočet hodnot daného rozdělení](#).

Kde všude se dá zadat vzorec

Vzorce (funkce) pro tvorbu či změnu proměnných se dají zadat na několika místech.

1. V dialogu proměnné v okénku pro dlouhý název
2. Ve výrazech pro dávkovou transformaci
3. Změnou dlouhého jména ve specifikaci všech proměnných

The image shows three overlapping dialog boxes from the Statistica software. The top box is 'Dlouhé jméno (popis či výraz s funkcemi):' with the formula '=iif(v1>v2;1;0)'. The middle box is 'Výrazy pro dávkovou transformaci' with the formula 'v3=iif(v1>v2;1;0)'. The bottom box is 'Editor specifikace proměnných' showing a table of variables.

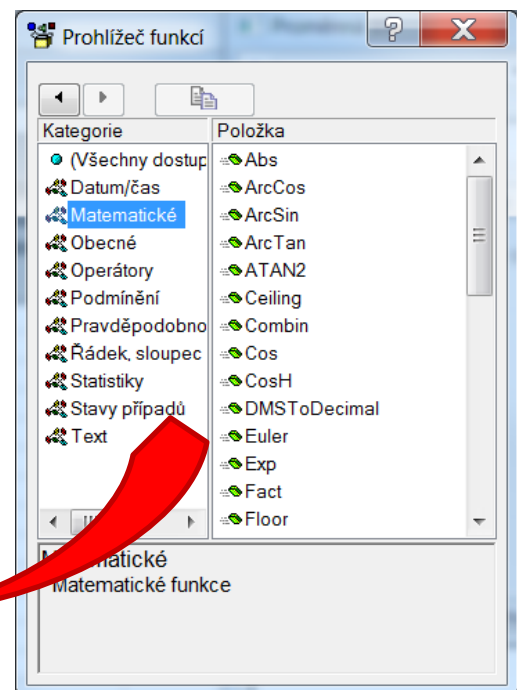
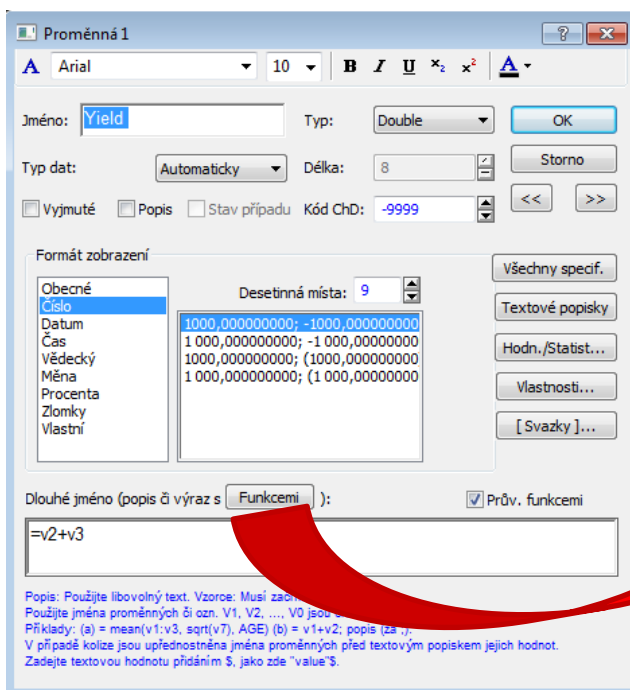
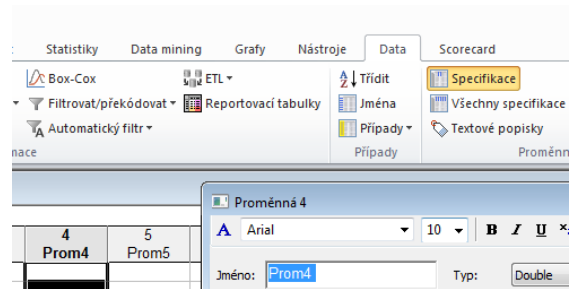
	Jméno	Typ	Kód ChD	Délka	Dl. jméno (popis, výraz, propojení)	Typ dat	Vyjmuté	F
1	Prom1	Double	-999999998			Autom	<input type="checkbox"/>	
2	Prom2	Double	-999999998			Autom	<input type="checkbox"/>	
3	Prom3	Double	-999999998			Autom	<input type="checkbox"/>	
4	přírůstek	Double	-999999998		=iif(v1>v2;1;0)	Autom	<input type="checkbox"/>	

U každého z nich si popíšeme výhody a nevýhody.

1. Specifikace proměnné

V dolní části dialogového okna *Specifikace proměnné* (vyvoláte jej například dvojklikem na záhlaví proměnné, kde chcete vložit vzorec, nebo přes záložku *Data*) je okénko pro „dlouhé jméno“, můžete zde tedy napsat nějaký komentář k proměnné nebo vzorec nebo obojí. Několik postřehů, jak zadávat vzorce:

- › Pokud se jedná o vzorec, musí začínat znakem rovníčka.
- › Pokud se odkazují na proměnnou, mohou buď celým jménem (pokud je v názvu mezera, pak navíc musím dát název do uvozovek) nebo číslem proměnné, pokud chci mít ve vzorci druhou proměnnou, pak napíšu v_2 (v z anglického *variable*).



- › Dostupné funkce vyvoláte tlačítkem *Funkcemi*.
- › Při psaní vzorců Vám pomáhá našeptávač, takže víte, co musíte napsat dále.
- › Argumenty funkcí jsou pro české nastavení oddělovačů desetinných míst odděleny středníkem.

Výhody

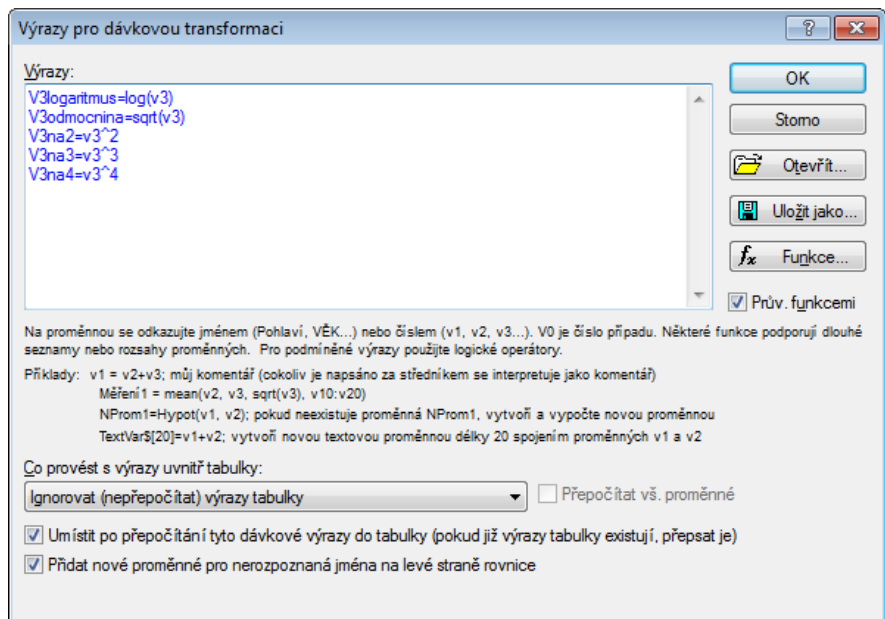
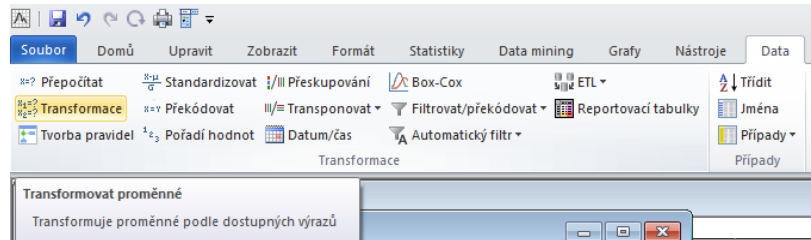
- › Jednoduchá definice dané (například nové) proměnné.
- › Vše okolo dané proměnné je možné nastavit v jednom dialogu.

2. Vzorce dávkových transformací

Tato možnost umožňuje definici hned několika vzorců najednou. V syntaxi je rozdíl pouze v tom, že před rovnítko je potřeba napsat proměnnou, kterou chceme měnit – zde totiž nemáme předem určeno, kterou proměnnou měníme, musíme to tedy specifikovat ve „zdrojovém kódu“. Možnost dávkových transformací najdete v záložce **Data** pod **Vzorce dávkových transformací** nebo v zobrazení typu pás karet jednoduše pod **Transformace**.

Pokud zadáme před rovnítko název proměnné, který neodpovídá žádnému názvu existujících proměnných, software vytvoří novou proměnnou na konci souboru s tímto názvem a podle nadefinovaného vzorce.

Vpravo vidíte příklad, kdy najednou nadefinujeme 5 proměnných určujících různé mocninné transformace proměnné číslo 3.

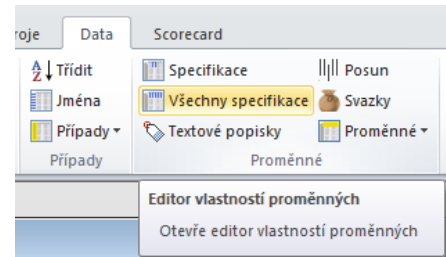


Výhody

- > Metoda je vhodná při definování více funkcí najednou.
- > Můžete takto přímo vytvořit velký počet nových proměnných.
- > Jednoduchá opakovatelnost. Pomocí tlačítek **Uložit jako** a **Otevřít** máte možnost uložit a poté zpětně vyvolat již jednou nadefinované příkazy.

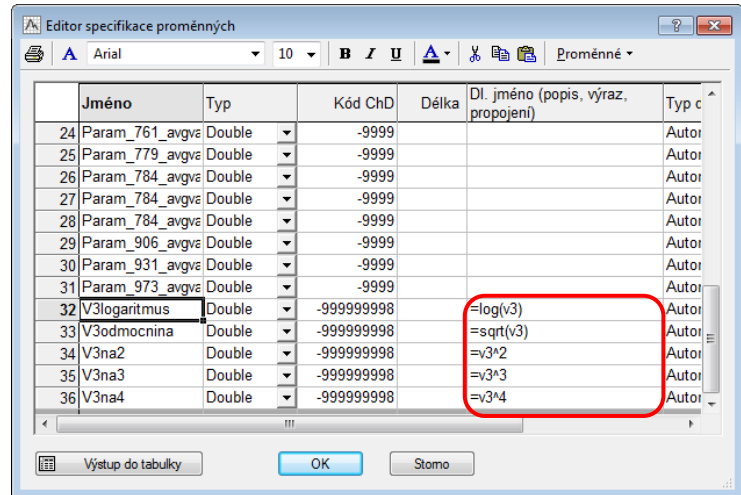
3. Všechny specifikace

Poslední možností je změnit dlouhé jméno v dialogu specifikací všech proměnných na příslušný vzorec.



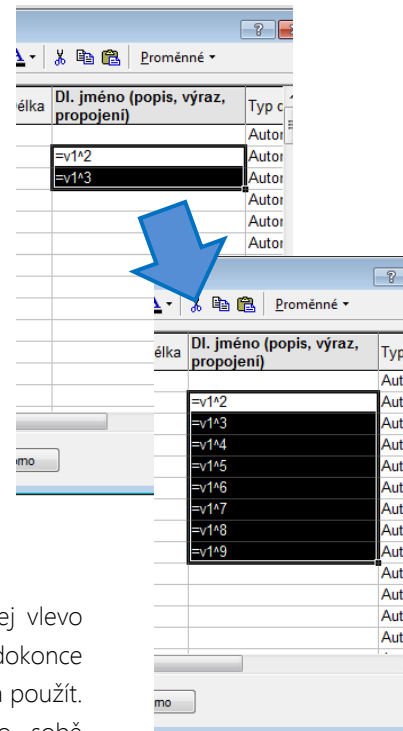
Výhody

- › Tato možnost je opět výhodná, pokud chcete nadefinovat více proměnných.
- › Všechny proměnné vidíte přehledně i s dalšími vlastnostmi proměnných.
- › Je zde možnost využít pro editaci vzorců specifikace práce s tabulkou, jelikož jsou vzorce v klasické tabulce – můžete tedy použít kopírování buněk, či vkládání buněk z jiného umístění.



Příklad 1: Využití kopírování.

Označíte buňky se vzorci, v pravém dolním rohu označovacího rámečku se objeví +, přetáhnete vzorec do dalších řádků a máte vzorce vyplněny rychleji a jednodušeji, než kdybyste to dělali ručně.

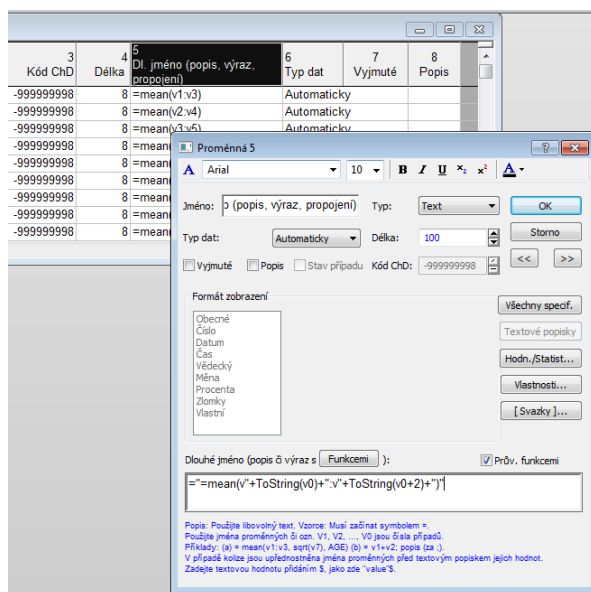


Příklad 2: Přenesení vzorců z jiné tabulky

Představte si případ, kdy máte v jedné tabulce již nadefinované vzorce a chcete ty stejné vzorce přenést do nové tabulky – můžete to provést jednoduše tak, že překopírujete vzorce ze specifikací proměnných jedné tabulky do specifikací proměnných tabulky druhé.

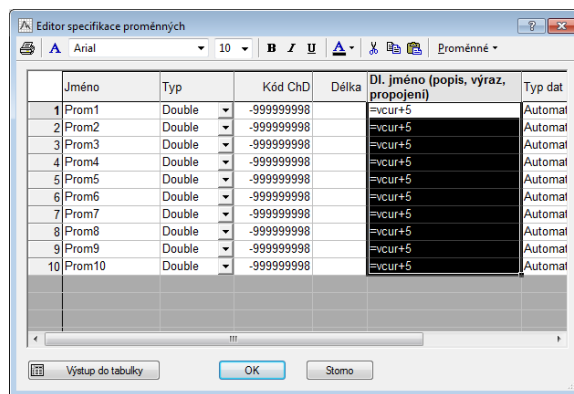
Příklad 3: Nastavení vzorce pro vzorce

Pokud využijete tlačítko **Výstup do tabulky** (najdete jej vlevo dole v editoru specifikace proměnných), můžete si dokonce nějaké složitější funkce naprogramovat a poté přenést a použít. Například chceme vytvořit klouzavý průměr tří po sobě jdoucích proměnných.



Tedy potřebujeme postupně vzorce: $= (v1+v2+v3)/3$, $= (v2+v3+v4)/3$,... nebo jinak řečeno: $= mean(v1:v3)$, $= mean(v2:v4)$,... Vkládání vzorců ručně jeden po druhém by u velkých datových souborů zabralo mnoho času. Udělejme to tedy tak, že si výrazy jednotlivých průměrů vytvoříme vzorcem. Tento vzorec musí vytvořit texty, které budou přesnými příkazy, které bychom rádi dále použili. Řešení je vlevo v obrázku. Funkce *ToString* převádí číslo na text, operace $+$ spojuje dva texty před a za znaménkem.

Poznámka: Zajímavou funkcí, která může být výhodně využita při vkládání více vzorců najednou, je funkce *VCUR*. Ta vrátí hodnoty aktuální proměnné. Pokud chcete u každé proměnné přičíst ke stávajícím hodnotám například číslo 5, pak stačí vložit do všech řádků vzorec $=vcurl+5$.



Důležité poznámky na konec

Jak už bylo řečeno, vzorce se zadávají jako sloupcové, tedy pro všechny buňky jednoho sloupce najednou.

Funkce nejsou automaticky přepočítávány po změně dat! Software můžete nicméně donutit přepočítat výrazy pomocí: *Data-přepočítat výrazy tabulky...* Dokud nezměníte vzorec nebo nedáte přepočítat, zůstanou tedy hodnoty v buňkách takové, jaké byly před změnami.