



Svazky proměnných

Newsletter Statistica ACADEMY



Téma: Možnosti softwaru
Typ článku: Tipy a triky, návod

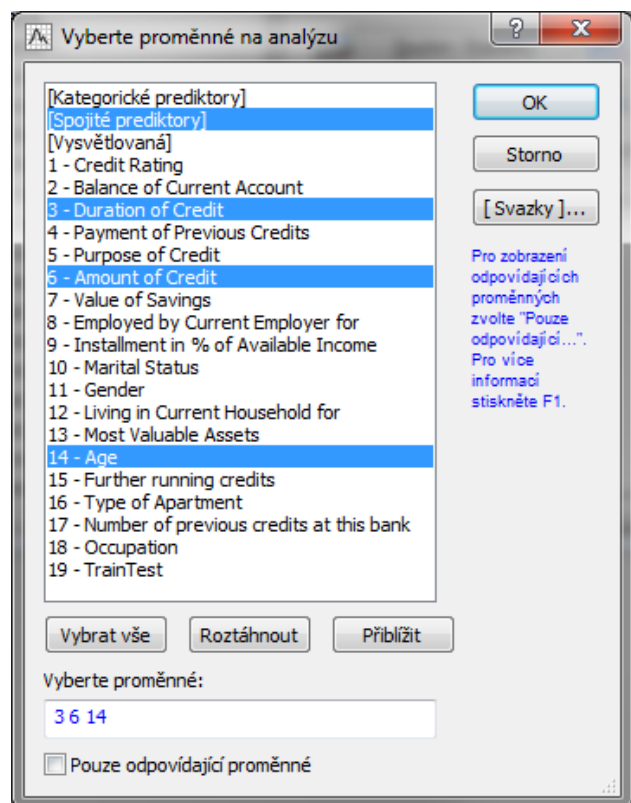
Ukážeme si, jak může využít při práci v softwaru funkcionalitu Svazky.

O co jde?

Svazky (v anglické verzi nazývané Bundles) jsou jakýmsi uměle vytvořenými seznamy proměnných. Pokud jsou v tabulce definovány, pak je možné je vidět a zvolit kdekoli, kde si lze volit proměnné. Volba svazku vždy vystupuje jako výběr několika proměnných najednou (což můžete vidět například na obrázku vpravo, místo na tři různé proměnné ve výčtu jsme tedy pouze klikli na svazek nazvaný Spojité prediktory). Pokud naopak vyberete proměnné a nějaký svazek je obsahuje, také se označí na znamení, že všechny proměnné, které obsahuje jsou v něm obsaženy.

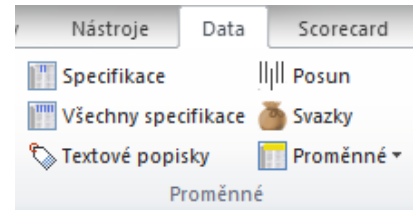
Svazky poznáte ve výčtu proměnných tak, že jsou v hranatých závorkách.

Svazky jsou uloženy spolu s tabulkou ve formátu .sta, o vytvořené svazky tedy nepřijdete a jsou jedním z atributů formátu tabulky Statistica.



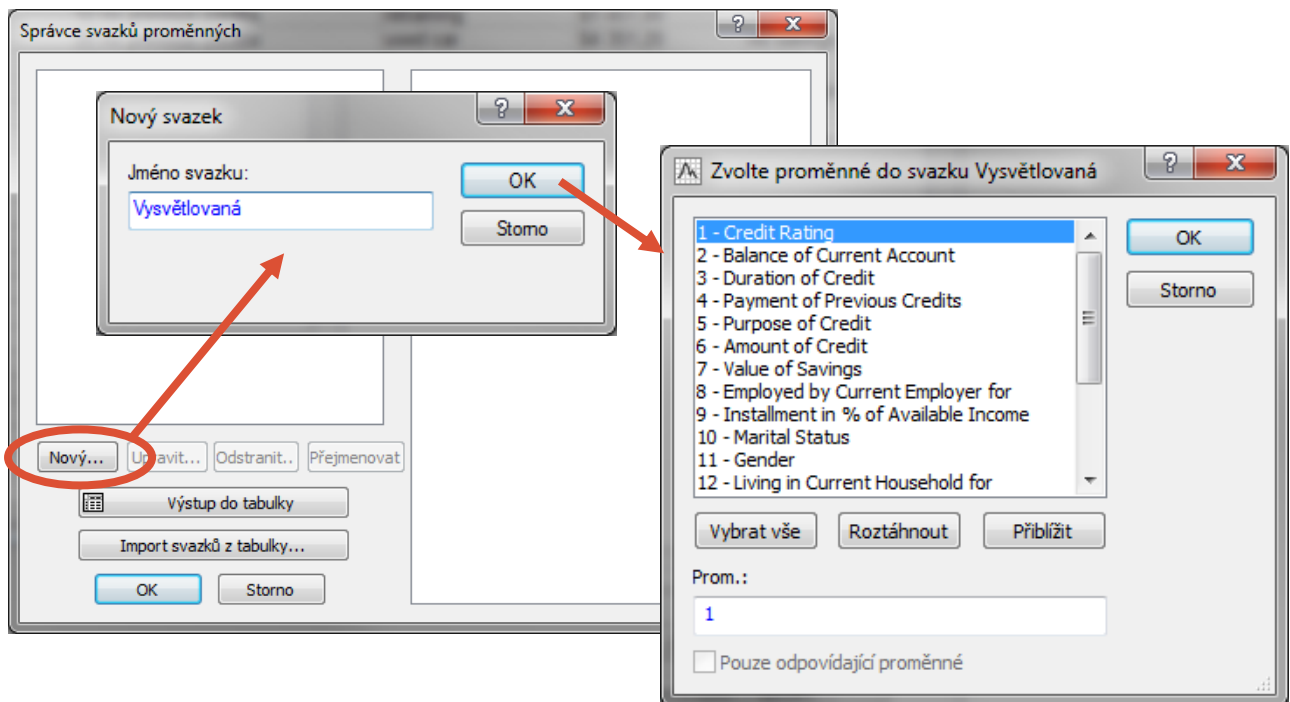
Jak nadefinovat svazky

Pokud svazky vytváříte poprvé, pak je se stačí vyvolat dialog **Správce svazků proměnných** a tam vše nadefinovat. Tento dialog se otevře, pokud klikneme na volbu **Svazky** v záložce Data, logicky se vyskytuje v podsekcí **Proměnné**.



Další možností je vyvolat tento dialog ve specifikacích kterékoli proměnné (tlačítko [**Svazky**]...).

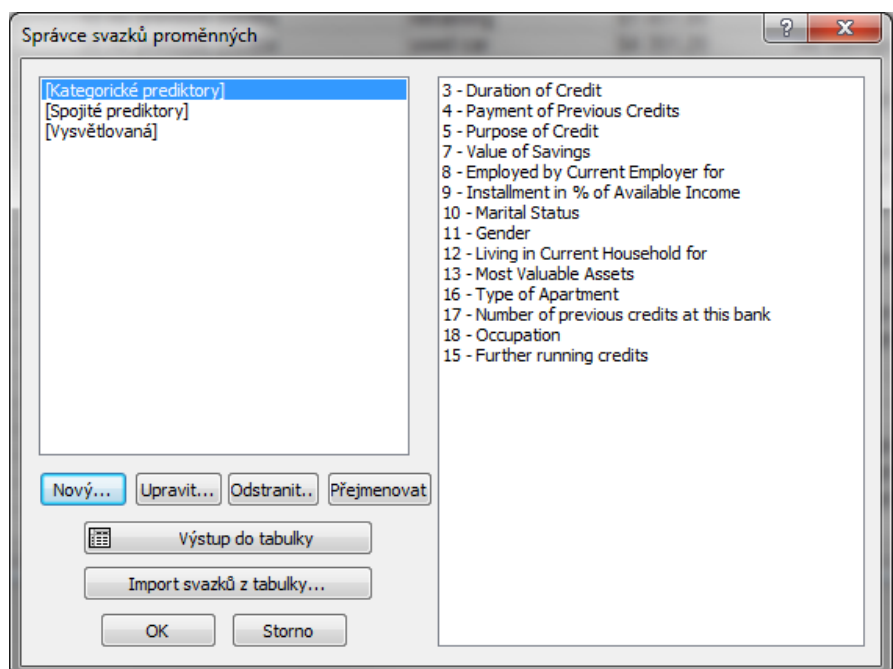
V dialogu **Správce svazků proměnných** pak postupně nadefinujeme to, co potřebujeme:



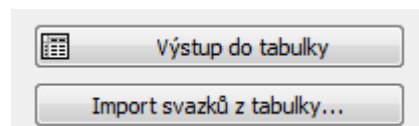
My jsme si postupně nadefinovali 3 svazky v souboru **CreditScoring.sta**. Jejich názvy jsou: **Vysvětlovaná**, **Kategorické prediktory**, **Spojité prediktory**.

Na obrázku vpravo vidíte výčet proměnných, které spadají do svazku Kategorické prediktory.

Svazky se mohou libovolně překrývat, jedná proměnná může být ve více svazcích. Prostě si nadefinujete přesně to, co se Vám bude hodit.



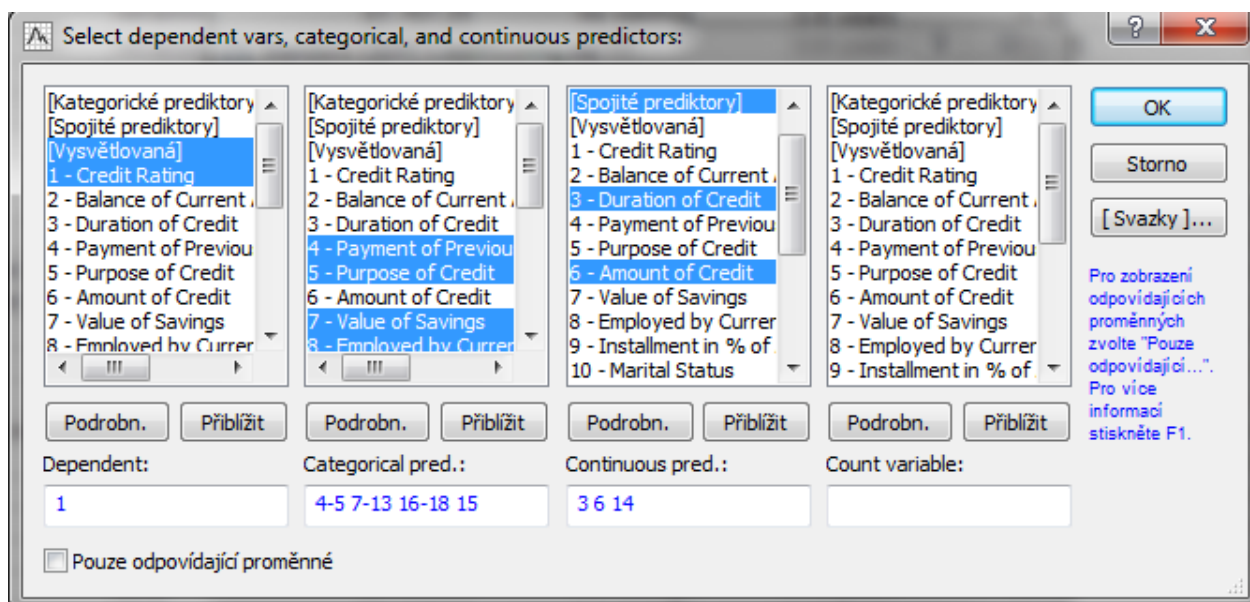
Druhou možností, jak svazky načíst je skrz již uložené svazky. Svazky lze načíst a také vyexportovat do tabulky pomocí tlačítek v dialogu, kde jsme svazky definovali: Pokud tedy máte svazky nadefinované již v jiném souboru, můžete je pomocí těchto tlačítek přenést.



K čemu jsou svazky

Jak jste mohli vidět už v prvním odstavci, pokud při volbě proměnných zvolíte nějaký svazek, automaticky se vyberou všechny proměnné, které svazek obsahuje. Pokud tedy máte hodně proměnných nebo chcete vybrat jen konkrétní skupinu proměnných, pak svazky jistě ulehčí práci.

Pokud si vezmeme naše svazky a budeme chtít vytvořit například model, kde Credit Rating vysvětlujeme pomocí kategorických a spojitých prediktorů, pak nám stačí pro výběr proměnných pouze 3 kliknutí, zatímco jednotlivých vybíráním bychom museli kliknout alespoň 8x:



Definováním svazků se tedy také vyhýbáme chybám a překlíkům při výběru proměnných (především v rozsáhlých a složitých datech).

Stejně tak svazky využijeme, pokud výběr stejných proměnných děláme opakovaně - ušetří nám to nejen čas, ale také to snižuje riziko, že bychom v některém z našich postupů zadali proměnné, které nechceme (například překlíkem).

Závěrem

Pokud se někdy dostanete do situace, kdy Vás začne otravovat, že musíte vybírat stejné proměnné pořád dokola, pak si vzpomeňte na svazky, V programu jsou právě proto, aby Vám v těchto situacích pomohly.