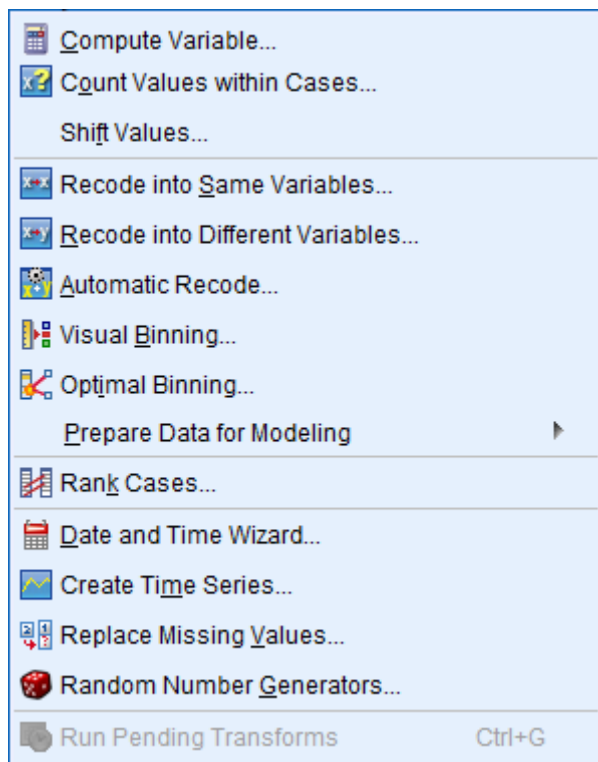


## 6 Transformace (Transform)



Nabídka Transform je určena především k odvozování nových proměnných nebo k jejich transformacím. K dispozici jsou tyto možnosti: výpočet nové proměnné na základě existujících (**Compute Variable**), odvození proměnné, která zjišťuje počet výskytů dané hodnoty na skupině proměnných (**Count Values within Cases**), posun hodnot (**Shift Values**), překódování hodnot proměnné do téže proměnné (**Recode into Same Variables**), překódování proměnné do nové proměnné (**Recode into Different Variables**), automatické překódování (**Automatic Recode**), vizuální kategorizace proměnných (**Visual Binning**), optimální kategorizace (**Optimal Binning**), příprava dat pro modelování (**Prepare Data for Modeling**), výpočet pořadových statistik (**Rank Cases**), průvodce pro práci s datem a časem (**Date and Time Wizard**), vytváření odvozených časových řad (**Create Time Series**) a nahrazení chybějících hodnot průměrem souboru nebo vybranou funkcí okolních bodů v časové řadě (**Replace Missing**

**Values**). Volba **Random Number Generators** je určena k nastavení generátoru náhodných čísel. Příkaz **Run Pending Transforms** provede zadané transformace, jejichž realizace byla odložena.

### Slovník anglických pojmů

<i>automatic</i>	.....	automatický
<i>case</i>	.....	případ
<i>compute</i>	.....	vypočítat
<i>count</i>	.....	počet
<i>different</i>	.....	jiný, odlišný
<i>missing</i>	.....	chybějící
<i>number</i>	.....	číslo
<i>random</i>	.....	náhodný
<i>rank</i>	.....	pořadí
<i>recode</i>	.....	překódování
<i>replace</i>	.....	přepsat, nahradit
<i>same</i>	.....	stejný, totožný
<i>value</i>	.....	hodnota
<i>variable</i>	.....	proměnná

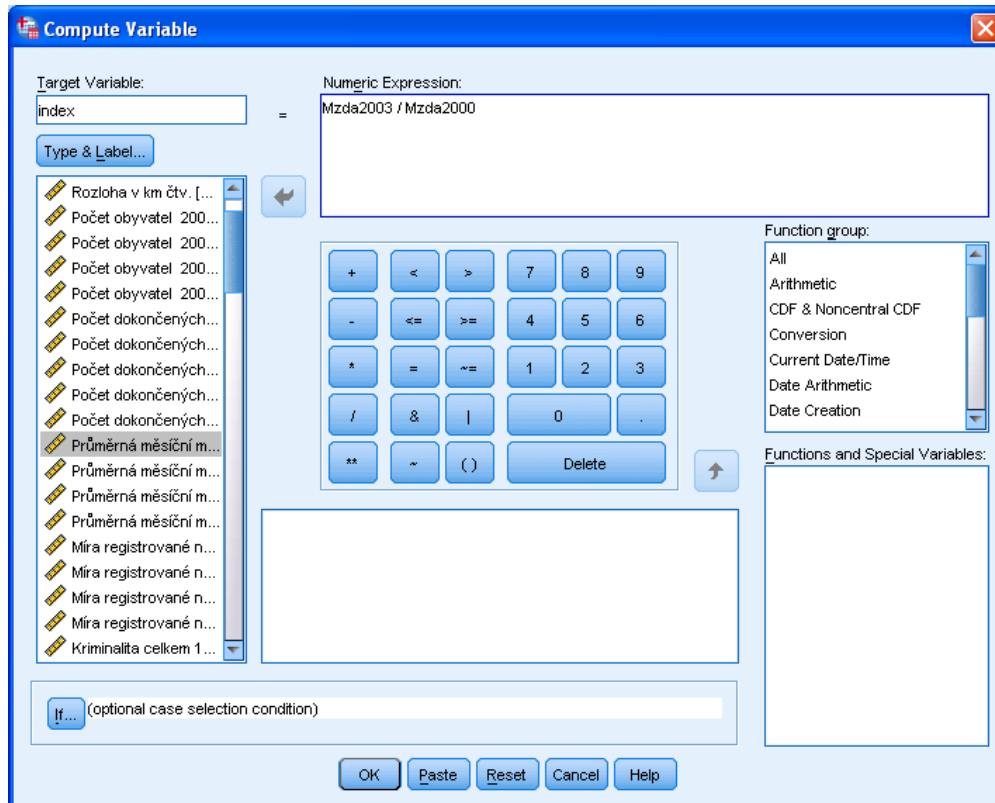
## 6.1 Výpočet nové proměnné (Compute Variable)

Procedura je určena k dovození nové proměnné na základě funkce existujících proměnných. K dispozici je celá řada funkcí od aritmetických a statistických až po funkce pro práci s textem, datem a časem a další speciální možnosti.

V poli **Target Variable** určíme název nové proměnné a případně jí přiřadíme také typ a popis (**Type & Label**). V poli **Numeric Expression** specifikujeme požadovaný výraz. Početní operace zadáme pomocí kalkulačky nebo přímo z klávesnice. Proměnné přenášíme ze seznamu v levé části okna šipkou nebo dvojitým poklepáním. Složitější funkce volíme v polích **Function group** a **Functions and Special Variables**. Transformační vzorec lze rovněž zapsat nebo upravit ručně.

V části **Function group** vyhledáváme požadované funkce podle skupin: vše (**All**), aritmetické (**Arithmetic**), kumulativní distribuční funkce (**CDF & Noncentral CDF**), konverzní (**Conversion**), aktuální datum a čas (**Current Date/Time**), aritmetické operace s daty (**Data Arithmetic**), vytvoření data (**Data Creation**), extrakce data (**Data Extraction**), inverzní distribuční funkce (**Inverse DF**), smíšené (**Miscellaneous**), chybějící hodnoty (**Missing Values**), hustoty rozdělení (**PDF & Noncentral PDF**), náhodná čísla (**Random Numbers**), vyhledávání (**Search**), significance (**Significance**), statistické (**Statistical**), funkce na skórování dat (**Scoring**), funkce pro práci s textovými řetězci (**String**), vytvoření časové proměnné (**Time Duration Creation**), extrakce času (**Time Duration Extraction**). Jakmile zvolíme jednu ze skupin, zobrazí se v poli **Functions and Special Variables** seznam dostupných funkcí. Po označení funkce se ve vedlejším okně objeví nápověda a funkci můžeme přenést do okna **Numeric Expression** šipkou nebo dvojitým poklepáním.

Dále můžeme upřesnit podmínku (**If**), kdy má být příkaz proveden. Není-li podmínka splněna, nebude hodnota cílové proměnné definována.



Výpočet nové proměnné

## Výpočet nové proměnné

- Transform
  - Compute Variable

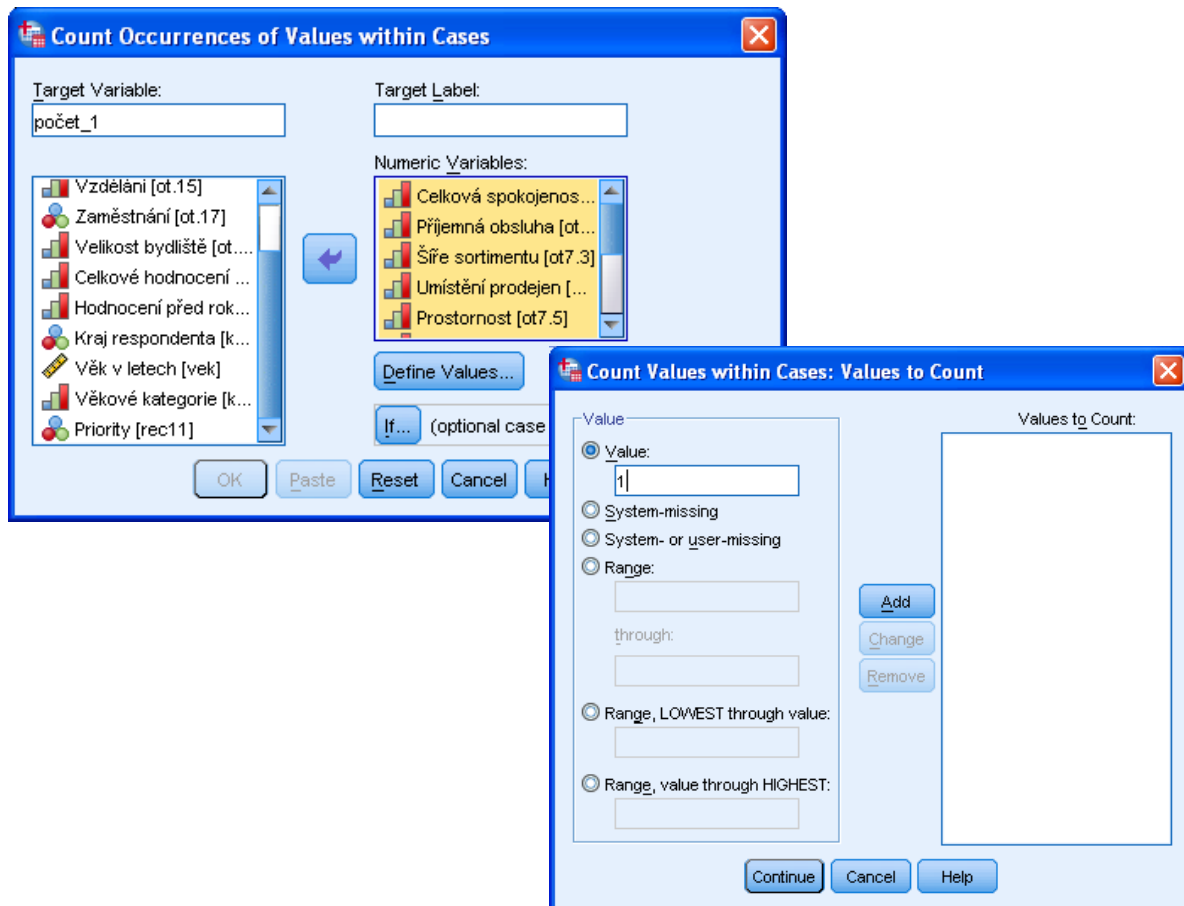
## 6.2 Počet výskytů hodnoty u případu (Count Values within Cases)

Procedura vytvoří novou proměnnou vyjadřující, kolikrát se vyskytnou vybrané hodnoty v daném případě na množině zvolených proměnných.

V poli **Target Variable** zadáme název nové proměnné, vyjadřující počet výskytů určitých hodnot u případu. Této proměnné můžeme rovněž přiřadit popis (**Target Label**). Do pole **Variables** přeneseme proměnné, u kterých budeme výskyt hodnot sledovat.

Tlačítkem **Define Values** definujeme zjišťované hodnoty. Lze zadat jednotlivé hodnoty (**Value**), systémové chybějící hodnoty (**System-missing**), systémové nebo uživatelem definované chybějící hodnoty (**System- or user-missing**), interval (**Range**), interval od mínus nekonečna do určité hodnoty (**Range, LOWEST through value**) nebo interval od určité hodnoty do nekonečna (**Range, value through HIGHEST**).

Podmínka **If** slouží pro výběr zpracovávaných případů. Jestliže není splněna, hodnota cílové proměnné není definována. (Funkce pro zadání podmínky viz oddíl 6.1 *Výpočet nové proměnné (Compute Variable)*, str. 135.)



**Určení počtu výskytů hodnot**

### Výpočet počtu výskytů

- Transform
  - Count Values within Cases

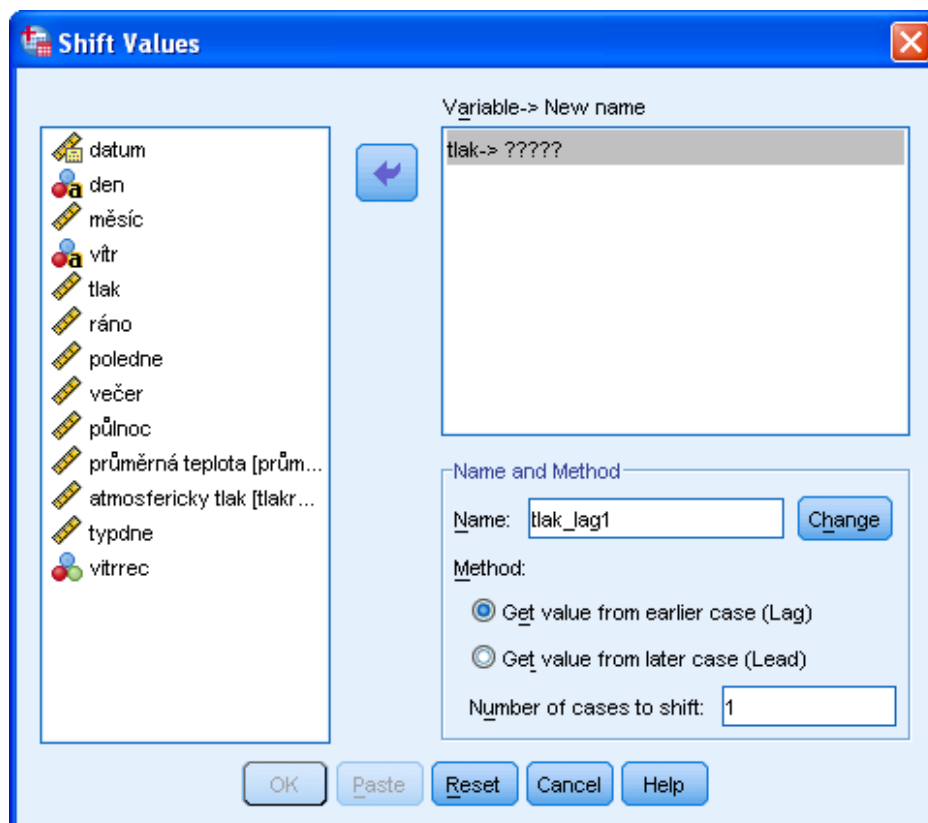
## 6.3 Posunutí hodnot proměnné (Shift Values)

Procedura vytvoří novou proměnnou, která obsahuje hodnoty existující proměnné posunuté o požadovaný počet míst dopředu (*lag*) nebo dozadu (*lead*). Tímto způsobem lze odvodit i více proměnných současně.

Proměnnou, ze které při odvozování vycházíme, přeneseme do okna **Variable** → **New name**. V poli **Name** zadáme název nové proměnné. V části **Method** definujeme způsob odvození proměnné:

- **Get value from earlier case (lag)**. Posunutí dopředu o požadovaný počet míst. Hodnoty předchozích případů z aktivního datového souboru (například při defaultním nastavení 1 bude každému případu přiřazena hodnota předchozího případu původní proměnné).
- **Get value from later case (lead)**. Posunutí dozadu o požadovaný počet míst. Hodnoty následujících případů z aktivního datového souboru (například při defaultním nastavení 1 bude každému případu přiřazena hodnota následujícího případu původní proměnné).

V poli **Number of cases to shift** určíme počet míst (přirozené číslo), o která se hodnoty posunou. Tlačítkem **Change** zadání potvrdíme.



Posunutí hodnot proměnné

### Posunutí hodnot proměnné

- Transform
  - Shift Values

## 6.4 Překódování hodnot do stejné proměnné (Recode into Same Variables)

Proceduru využijeme, jestliže chceme změnit systém kódování proměnné (např. spojit kategorie, změnit pořadí hodnot, definovat intervaly atd.). Při použití této nabídky program přepíše původní proměnnou, pro překódování do nové proměnné je určena nabídka **Recode into Different Variables**.

Nejprve vybereme proměnnou (proměnné), která se má překódovat a šipkou ji převedeme do pole **Variables**. Tlačítko **Old and New Values** slouží k zadání převodního pravidla. Na levé straně vyznačíme vždy původní hodnoty (**Old Value**), na pravé straně nové hodnoty (**New Value**).

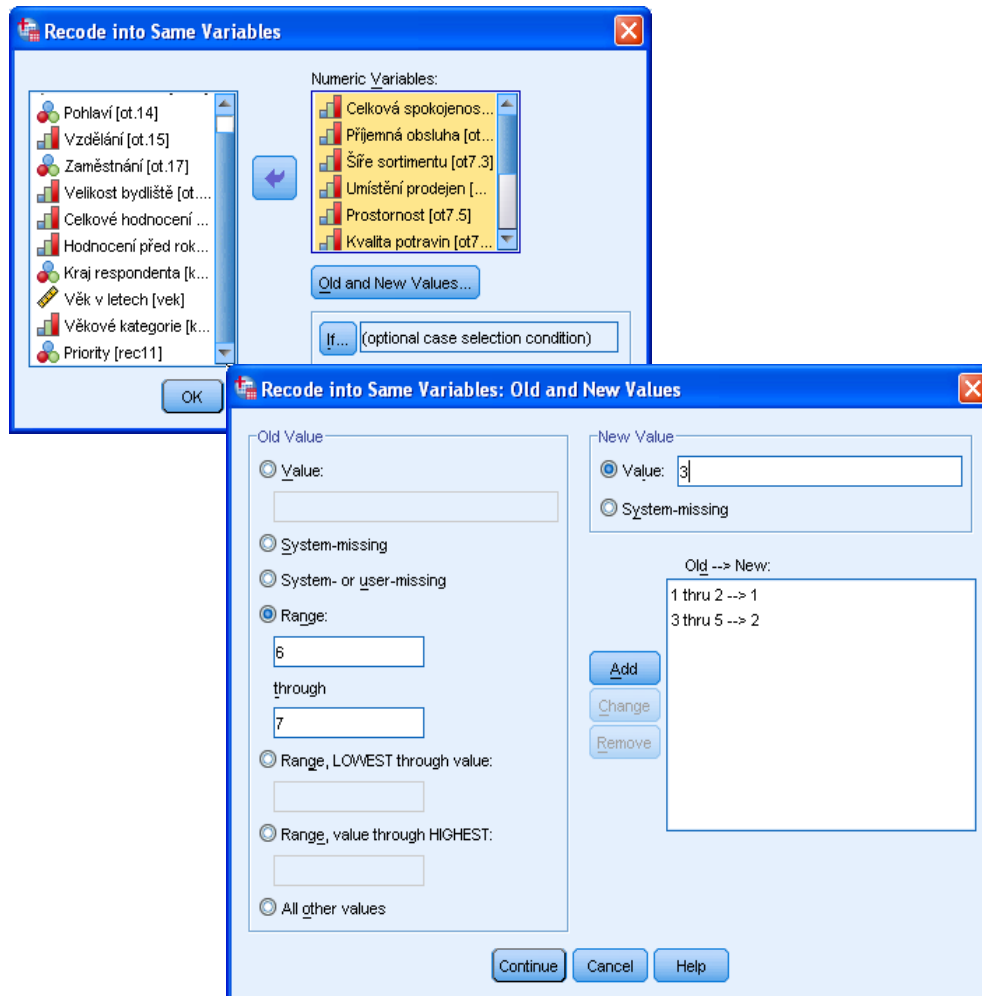
Původní hodnoty můžeme zadat jednu po druhé (**Value**), zvolit systémově vynechané hodnoty (**System-missing**), zvolit všechny hodnoty definované jako vynechané (**System- or user-missing**), vybrat hodnoty uvnitř intervalu (**Range**), interval od mínus nekonečna do určité hodnoty (**Range, LOWEST through value**) nebo interval od určité hodnoty do nekonečna (**Range, value through HIGHEST**) nebo změnit všechny ostatní nespecifikované hodnoty hromadně (**All other values**).

Nové hodnoty lze zadat postupně (**Value**) nebo jako systémově chybějící (**System-missing**).

Položku je třeba vždy přidat do seznamu **Old** → **New** pomocí tlačítka **Add**.

Původní hodnoty, které nejsou převodním pravidlem změněny, zůstávají stejné.

Tlačítkem **If** můžeme dále zadat podmínku pro výběr zpracovávaných případů. Jestliže není splněna, hodnota cílové proměnné není definována. (Funkce pro zadání podmínky viz oddíl 6.1 *Výpočet nové proměnné (Compute Variable)*, str. 135.)



Překódování hodnot do stejné proměnné

### Překódování hodnot do stejné proměnné

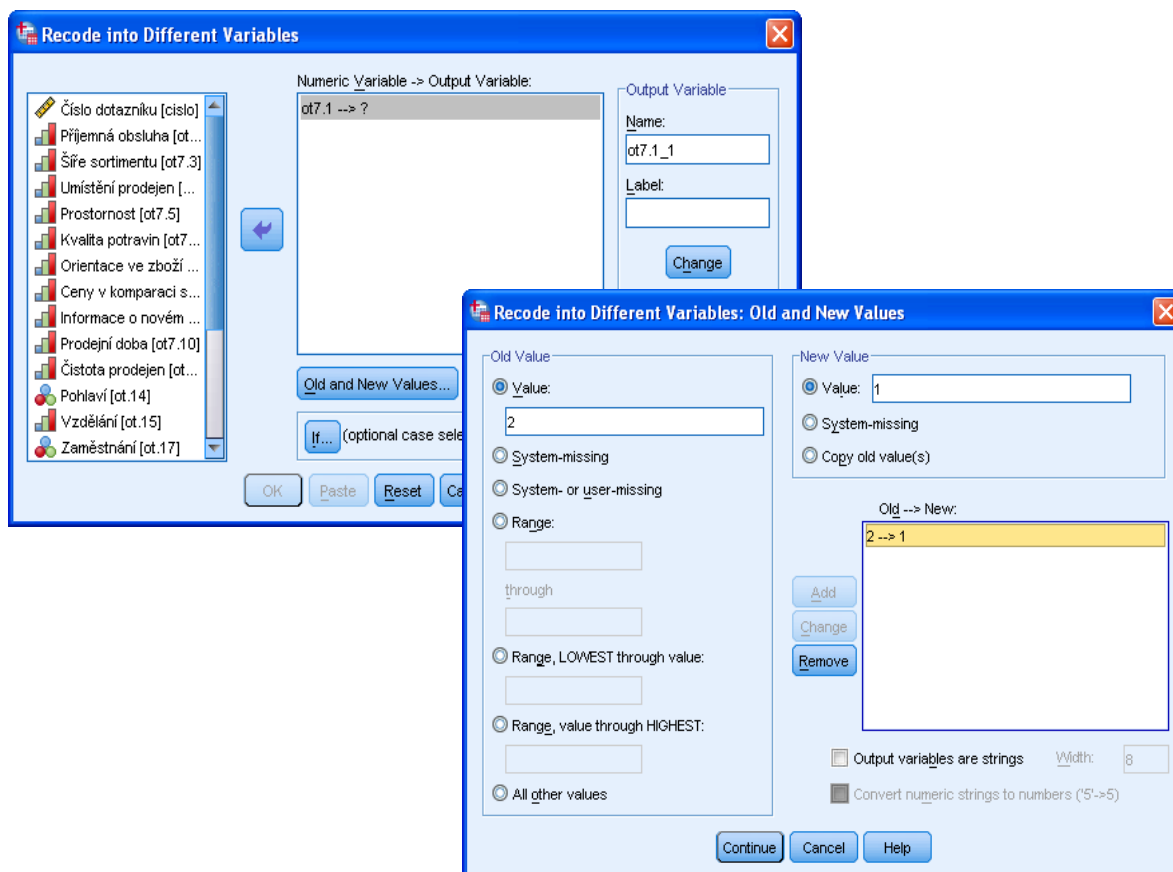
- Transform
  - Recode into Same Variables

## 6.5 Překódování hodnot do jiné proměnné (Recode into Different Variables)

Proceduru využijeme, jestliže chceme změnit systém kódování proměnné (např. spojit kategorie, změnit pořadí hodnot, definovat intervaly atd.). Při použití této nabídky program vytvoří novou proměnnou, pro přepsání původní proměnné je určena nabídka **Recode into Same Variables**.

Nejprve vybereme proměnnou (proměnné), kterou chceme překódovat a šipkou ji převedeme do pole **Input Variable** → **Output Variable**. Dále určíme jméno nové proměnné a případně její popis (**Output Variable**) a potvrdíme tlačítkem **Change**.

Další zadání příkazu je stejné jako v případě překódování hodnot do téže proměnné (viz oddíl 6.4 *Překódování hodnot do stejné proměnné (Recode into Same Variables)*, str. 137). Navíc je možné v okně **Old and New values** zadat, aby nová proměnná byla znaková (**Output variables are strings**) či změnit znakové symboly na čísla (**Convert numeric strings to numbers ('5' -> 5)**).



*Překódování do nové proměnné*

### **Překódování hodnot do jiné proměnné**

- *Transform*
  - *Recode into Different Variables*



**Slovník anglických pojmů**

<i>add</i> .....	<i>přidat</i>
<i>all</i> .....	<i>vše</i>
<i>condition</i> .....	<i>podmínka</i>
<i>convert</i> .....	<i>převést</i>
<i>copy</i> .....	<i>kopírovat</i>
<i>different</i> .....	<i>jiný, odlišný</i>
<i>change</i> .....	<i>změnit</i>
<i>include</i> .....	<i>zahrnout</i>
<i>input</i> .....	<i>vstupní</i>
<i>label</i> .....	<i>označení, popis</i>
<i>missing</i> .....	<i>chybějící</i>
<i>name</i> .....	<i>jméno</i>
<i>new</i> .....	<i>nový</i>
<i>numeric</i> .....	<i>číselný</i>
<i>old</i> .....	<i>starý</i>
<i>other</i> .....	<i>ostatní</i>
<i>output</i> .....	<i>výstup, výstupní</i>
<i>range</i> .....	<i>interval, rozpětí</i>
<i>recode</i> .....	<i>překódování</i>
<i>remove</i> .....	<i>zrušit</i>
<i>same</i> .....	<i>tentýž</i>
<i>satisfy</i> .....	<i>splňovat</i>
<i>string</i> .....	<i>znakový</i>
<i>system-missing</i> .....	<i>systémově vynechávaný</i>
<i>user-missing</i> .....	<i>uživatelsky vynechávaný</i>
<i>value</i> .....	<i>hodnota</i>
<i>variable</i> .....	<i>proměnná</i>

## 6.6 Automatické překódování (Automatic Recode)

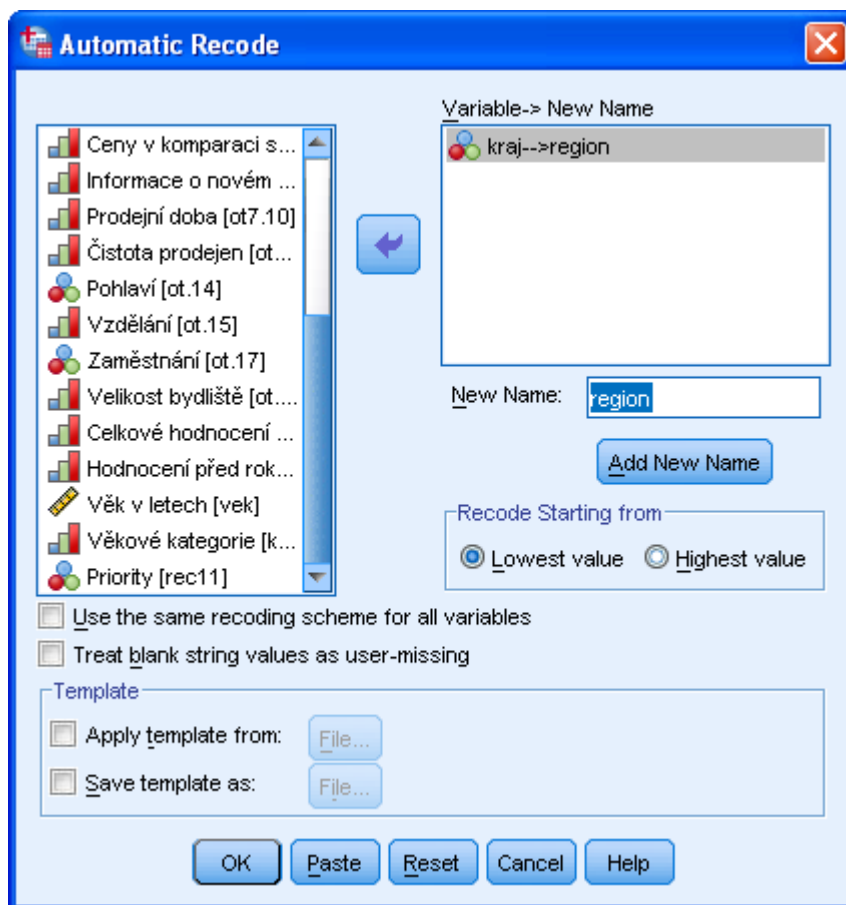
Procedura automaticky překóduje hodnoty proměnné do uspořádané řady celých čísel od 1 výše. Původní kategorie mohou být překódovány od nejvyšší do nejnižší nebo naopak. Jestliže se jedná o textovou proměnnou, vychází se z uspořádání podle abecedy.

Do pole **Variable** → **New Name** přeneseme proměnnou, jejíž hodnoty chceme překódovat. V dalším poli zadáme název nové proměnné a potvrdíme stisknutím tlačítka **Add New Name**. Na závěr zvolíme, zda chceme začít překódování od nejnižší nebo od nejvyšší hodnoty (**Recode Starting from**).

V dolní části okna jsou k dispozici volby:

- **Use the same recoding scheme for all variables**  
Pokud je stejný textový řetězec obsažen ve více proměnných, je mu přiřazen vždy stejný kód.
- **Treat blank string values as user-missing**  
Prázdné hodnoty textového řetězce jsou považovány za uživatelem definované chybějící hodnoty a je jim přiřazen nejvyšší kód.

Procedura dovoluje uložit šablonu určitého překódování (**Save template as**). Pokud se při dalším překódování šablona načte (**Apply template from**), přiřadí se kódy ze šablony těm textovým řetězcům v nové proměnné, které jsou shodné s řetězcem uloženými v šabloně.



Automatické překódování

### Automatické překódování

- Transform
  - Automatic Recode

## 6.7 Vizuální kategorizace (Visual Binning)

**Visual Binning** umožňuje snáze vytvářet nové proměnné založené na konečném počtu nepřekrývajících se skupin původní proměnné.

V prvním dialogovém okně pouze volíme proměnné, se kterými budeme pracovat. Pro případ, že pracujeme s velkým souborem, je rovněž možné omezit počet případů, které se při vizualizaci načítají.

Jakmile v následujícím okně označíme vybranou proměnnou, objeví se v pravé části název a popis této proměnné, informace o minimální a maximální hodnotě a její grafické znázornění. V části **Binned Variable** zadáme název a popis nově odvozené proměnné. V nabídce **Upper Endpoints** volíme typ nerovnosti určující, které ze dvou sousedních skupin budou náležet krajní body.

Tlačítkem **Make Cutpoints** zadáme způsob výpočtu intervalů, k dispozici jsou přitom tři základní volby:

- **Equal Width Intervals** – intervaly stejné šířky
- **Equal Percentiles Based on Scanned Cases** – stejné procento případů (založeno na načtených případech)
- **Cutpoints at Mean and Selected Standard Deviations Based on Scanned Cases** – zlomový bod se nachází v průměru a dále ve vzdálenostech odpovídajícím zvolenému násobku směrodatné odchylky (založeno na načtených případech)

Tam, kde je to nutné, upřesníme ještě polohu prvního zlomového bodu, počet zlomových bodů nebo šířku intervalů.

Tlačítkem **Make Labels** se automaticky vytvoří popisy hodnot vycházející z definice jednotlivých kategorií.

Zaškrtnutím políčka **Reverse scale** obrátíme číslování kategorií tak, že nejvyšší hodnota bude odpovídat nejnižším hodnotám původní proměnné.

Intervaly lze upravovat podle potřeby také ručně změnou hodnot v tabulce ve sloupečku **Value** nebo pomocí myši přesunutím oddělovače v grafu do jiné polohy.

Tlačítka **Copy Bins From Another Variable** a **To Other Variables** dovolují zkopírovat definice kategorií z jiné proměnné nebo na jinou proměnnou. U takto zadaných proměnných je však nutné ještě doplnit název, aby se nová proměnná skutečně vytvořila.

**Visual Binning**

Scanned Variable List: Věk v letech [vek]

Name: vek Label: Věk v letech

Current Variable: vek Binned Variable: věk\_kat Věk v letech (Binned)

Minimum: 18 Nonmissing Values Maximum: 90

Grid:

	Value	Label
1	30,0	<= 30
2	40,0	31 - 40
3	50,0	41 - 50
4	60,0	51 - 60
5	70,0	61 - 70
6	80,0	71 - 80
7		HIGH 81+
8		

Upper Endpoints  
 Included (<=)  
 Excluded (<)

Buttons: Make Cutpoints..., Make Labels, Reverse scale

Buttons: OK, Paste, Reset, Cancel, Help

Visual Binning

## Vizuální kategorizace

- Transform
  - Visual Binning

### Slovník anglických pojmů

<i>all</i>	vše
<i>case</i>	případ
<i>compute</i>	výpočet
<i>continue</i>	pokračovat
<i>count</i>	počet
<i>delete</i>	odstranit
<i>expression</i>	vyjádření; výraz
<i>function</i>	funkce
<i>high</i>	vysoký
<i>if</i>	jestliže
<i>include</i>	zahrnout
<i>label</i>	označení, popis
<i>low</i>	nízký
<i>missing</i>	chybějící
<i>numeric</i>	číselný
<i>range</i>	interval
<i>satisfy</i>	splňovat
<i>target</i>	cíl
<i>through</i>	až
<i>transform</i>	přeměnit
<i>type</i>	typ
<i>variable</i>	proměnná

## 6.8 Optimální kategorizace (Optimal Binning)

Procedura **Optimal Binning** je určena k optimální kategorizaci jedné nebo více číselných proměnných vzhledem k zadané kategorizované proměnné. Nově odvozenou kategorizovanou proměnnou (proměnné) lze následně užít pro další analýzu. Algoritmus je založen na metodě **MDLP (minimal description length principle)** a na statistice **Entropie**.

Na záložce **Variable** zadáme do pole **Variables to Bin** číselné proměnné, které mají být kategorizovány, a do pole **Optimize Bins with Respect To** kategorizovanou proměnnou, vzhledem k níž má být kategorizace optimální.

Na záložce Output označíme požadované výstupy:

- **Endpoints for bins** – nalezené optimální dělicí body pro každou kategorizovanou vstupní číselnou proměnnou a četnosti řídicí kategorizované proměnné ve skupinách,
- **Descriptive statistics for variables that are binned** – popisné statistiky vstupních číselných proměnných (počet, minimum, maximum, počet různých hodnot a počet nalezených dělicích bodů),
- **Model entropy for variables that are binned** – hodnota statistiky *Entropie* pro každou kategorizovanou vstupní číselnou proměnnou vzhledem k řídicí kategorizované proměnné.

Na záložce **Save** volíme, jaké informace mají být uloženy.

V části **Save Variables to Active Dataset** určíme, zda mají být kategorizované proměnné uloženy do datové matice (**Create variables that contain binned data values**) a zda v případě konfliktu názvů mají být nahrazeny dříve vytvořené proměnné (**Replace existing variables that have the same name**).

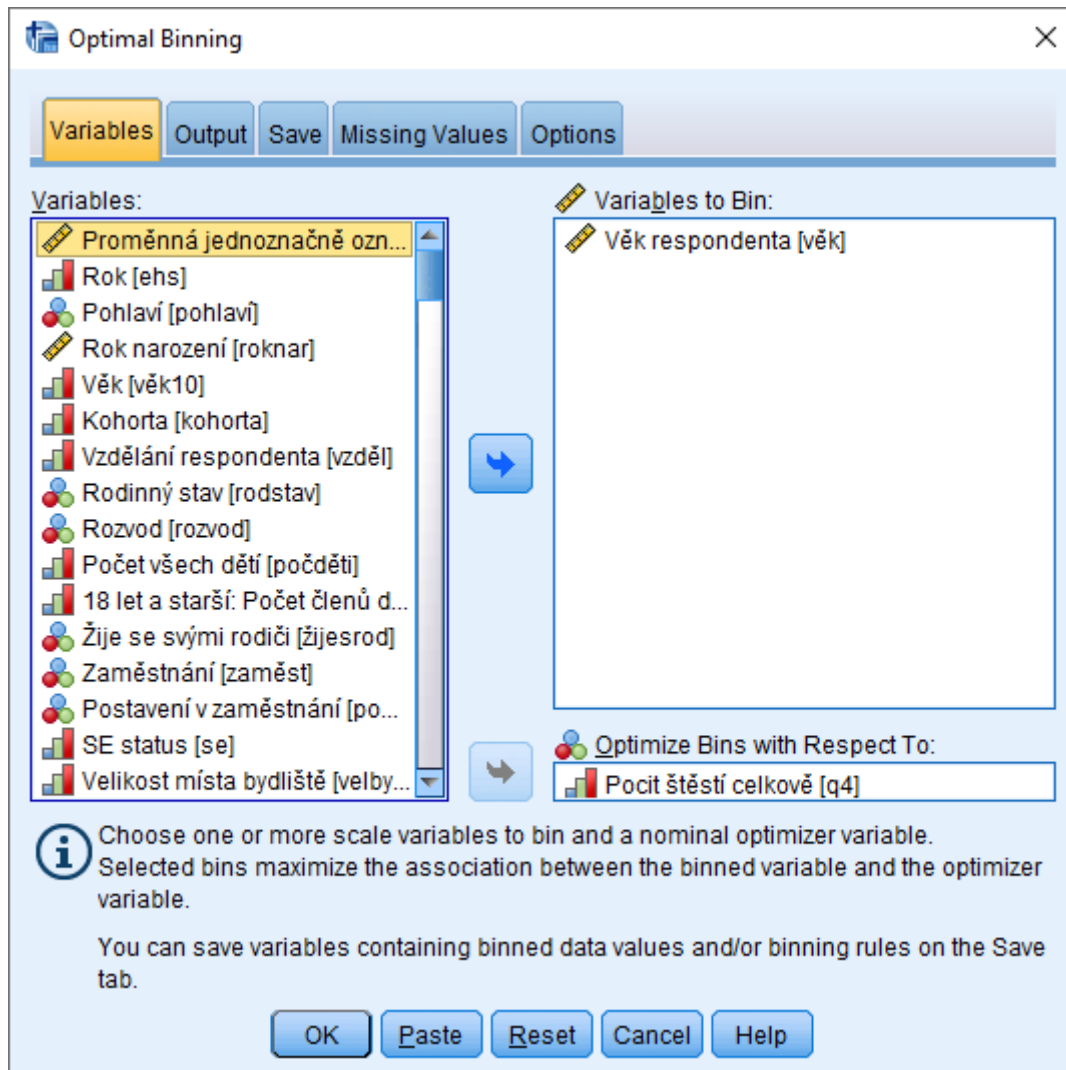
V části **Save Binning Rules as Syntax** lze zadat uložení pravidel pro odvození kategorizovaných proměnných do syntaxe.

Na záložce **Missing Values** volíme způsob práce s vynechanými hodnotami:

- **Pairwise** – pro každou dvojici (číselná proměnná a řídící kategorizovaná) samostatně,
- **Listwise** – pokud má kterákoliv ze vstupních proměnných vynechanou hodnotu, je případ vyloučen.

Uživatелеm definované vynechané hodnoty jsou vždy považované za neplatné a při kategorizaci jsou převedeny na systémové vynechané hodnoty.

Na záložce **Options** lze provést další nastavení, která se týkají urychlení algoritmu pro velké datové soubory (**Preprocessing**), spojení řídce zastoupených kategorií (**Sparsely Populated Bins**), rozhodnutí, do které z kategorií budou zahrnuty dělicí body (**Bin Endpoints**) a zda krajní kategorie budou od mínus nekonečna resp. do plus nekonečna nebo budou začínat nejnižší resp. končit nejvyšší hodnotou v datech.



Optimální kategorizace

### Optimální kategorizace

- Transform
  - Optimal Binning

## 6.9 Příprava dat pro modelování (Prepare Data for Modeling)

### 6.9.1 Interaktivní příprava dat (Interactive)

Procedura je určena k přípravě dat pro analýzu nebo modelování. Uživatel přitom může pracovat v poloautomatickém režimu nebo interaktivně volit, které kroky se budou realizovat.

Na záložce **Objective** určíme hlavní cíl přípravy dat:

- **Balance speed & accuracy** – rychlost při zpracování dat i přesnost modelu (obojí se stejnou váhou),
- **Optimize for speed** – optimalizace vzhledem k rychlosti zpracování dat,
- **Optimize for accuracy** – optimalizace s ohledem na přesnost modelu,
- **Custom analysis** – manuální nastavení algoritmů na záložce *Settings*.

Na záložce **Fields** specifikujeme vstupní proměnné. Ty je možné zadat manuálně (**Use custom field assignment**) nebo využít předdefinované role (**Use predefined roles**). Do pole **Target (optional)** lze volitelně zadat cílovou proměnnou. Ostatní vstupní proměnné přeneseme do pole **Inputs**.

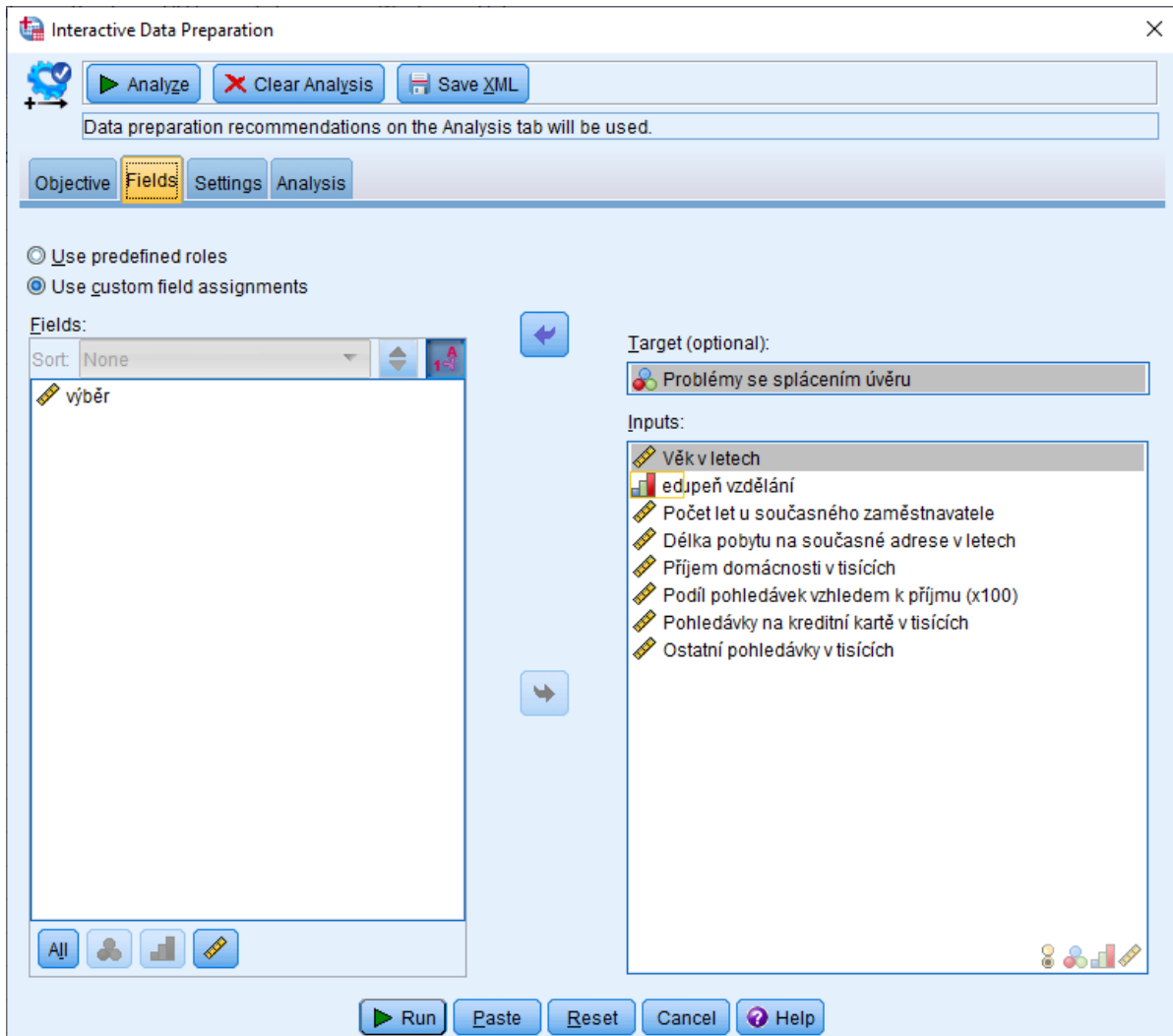
Záložka **Settings** umožňuje nastavit požadované kroky přípravy dat (pokud je na záložce *Objective* zvolena jiná možnost než *Custom analysis*, odpovídají zde přednastavené volby danému cíli). V části **Select an item** lze zadat následující:

- **Prepare dates & times** – příprava proměnných vyjadřujících datum a čas,
- **Exclude Fields** – vyloučení problematických proměnných (splňujících uvedená kritéria z analýzy),
- **Adjust Measurement** – úprava způsobu měření proměnné (číselné proměnné s malým počtem hodnot se změní na ordinální a ordinální proměnné s příliš mnoha různými hodnotami na číselné),
- **Improve Data Quality** – zlepšení kvality dat (nahrazení extrémních hodnot zvolenými hraničními hodnotami nebo vynechanými hodnotami, imputace vynechaných hodnot u různých typů proměnných, překódování nominálních proměnných od nejméně četné kategorie po nejčetnější)
- **Rescale Fields** – změna škály proměnných,
- **Transform Fields** – další typy transformací (slučování kategorií a hodnot),
- **Select And Construct** – nastavení hladiny významnosti pro výběr proměnných do modelu a odvození nových proměnných kombinací existujících,
- **Name Fields** – způsob odvozování názvů nových proměnných,
- **Apply Transformations** – přidání nových proměnných do aktivního datového souboru nebo vytvoření nového souboru.

Na záložce **Analysis** se po stisknutí tlačítka **Analyze** v hodní části okna zobrazují výsledky analýzy a informace o provedených akcích. Tento typ výstupu se chová obdobně jako výstupy do objektu *Model Viewer* (viz skripta *IBM SPSS Statistics Base 27, část Datový soubor a práce s výstupy*). Tlačítkem **Clear Analysis** tyto informace odstraníme.

Tlačítko **Save XML** v horní části okna umožňuje uložit transformace do formátu XML pro další využití (například pro zpětnou transformaci skóru).

Transformace proměnných se realizují až po spuštění procedury stiskem tlačítka **Run**.



*Interaktivní příprava dat*

### **Interaktivní příprava dat**

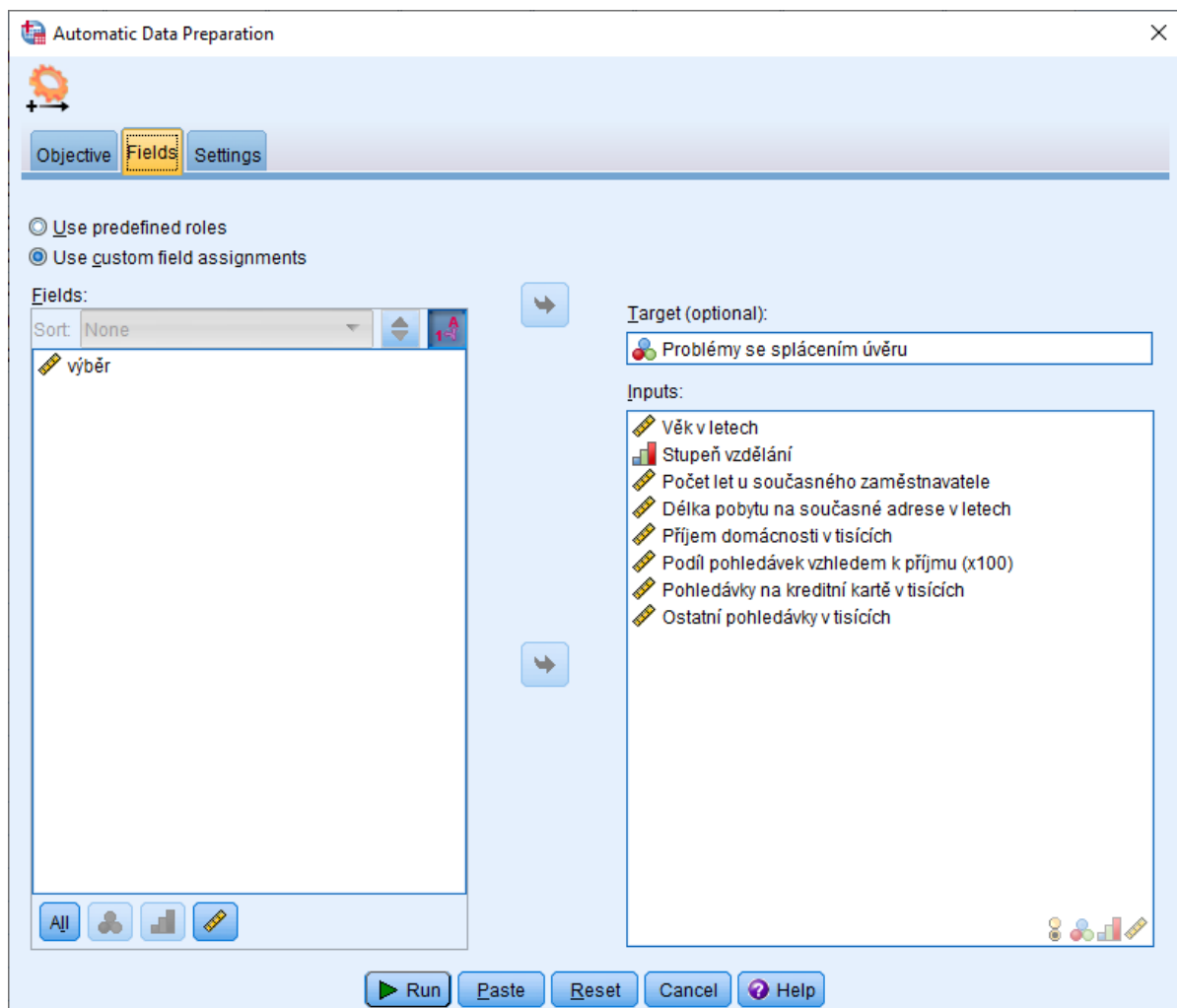
- *Transform*
  - *Prepare Data for Modeling*
    - *Interactive*



## 6.9.2 Automatická příprava dat (Automatic)

Procedura je určena k přípravě dat pro analýzu nebo modelování. Na rozdíl od předcházející procedury pro interaktivní přípravu dat je však zjednodušená, protože předpokládá, že uživatel bude pracovat především v automatickém režimu.

Způsob zadávání je obdobný jako v případě procedury **Interactive Data Preparation**, nejsou však k dispozici tlačítka *Analyze*, *Clear Analysis a Save XML* v horní části okna a chybí záložka *Analysis*. Naopak na záložce **Settings** je navíc možné při označení volby **Apply and Save** uložit transformace do syntaxového souboru (**Save transformations as syntax**) nebo do formátu XML (**Save transformations as XML**).



Automatická příprava dat

### Automatická příprava dat

- Transform

- Prepare Data for Modeling
  - Automatic

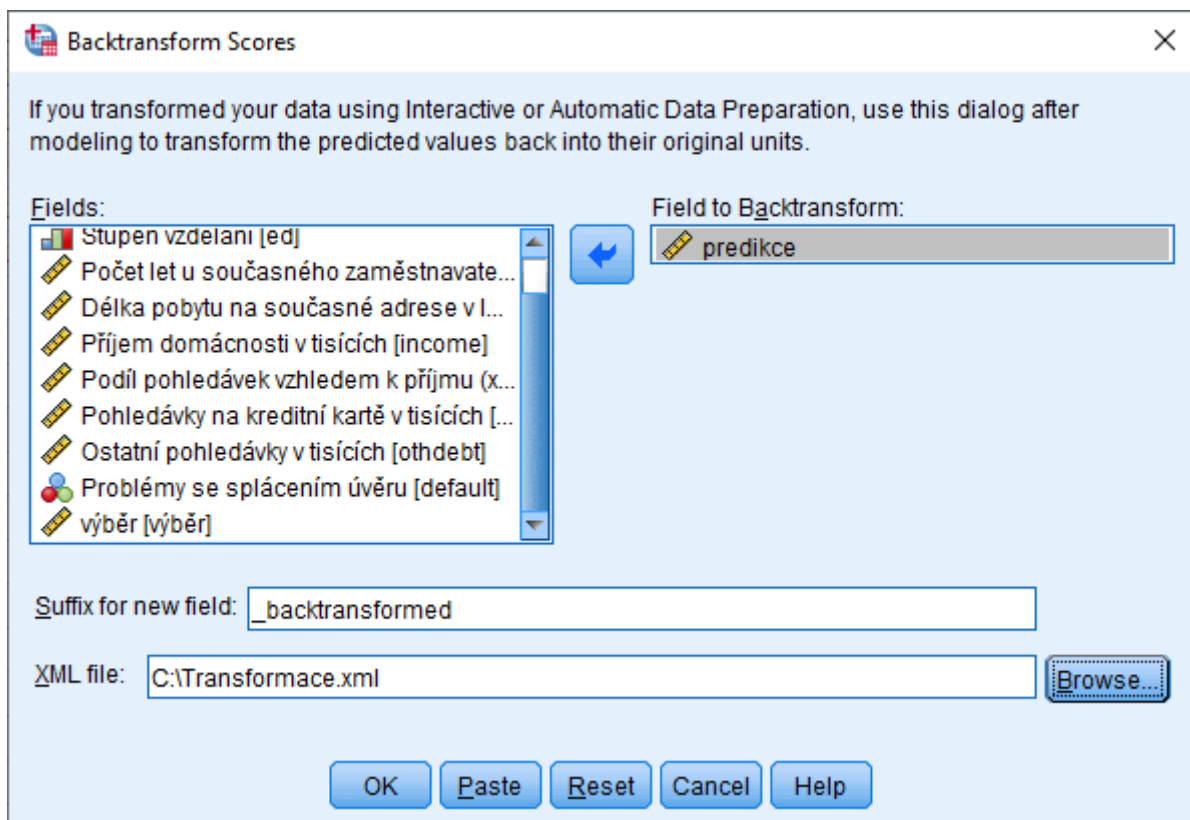
### 6.9.3 Zpětná transformace skóru (Backtransform Scores)

Proceduru je možné užít ke zpětné transformaci předpovědaných hodnot cílové proměnné na základě modelu v případě, že pro přípravu dat byly užity procedury *Interactive Data Preparation* nebo *Automatic Data Preparation*.

Do pole **Field to Backtransform** zadáme proměnnou obsahující předpovědané hodnoty cílové proměnné.

V poli **Suffix for new field** lze upřesnit příponu názvu nové proměnné.

V poli **XML file** specifikujeme XML soubor s transformacemi, který byl vytvořen procedurou *Interactive Data Preparation* nebo *Automatic Data Preparation*.



Zpětná transformace skóru

## Zpětná transformace skóru

- Transform
  - Prepare Data for Modeling
    - Backtransform Scores

## 6.10 Pořadové statistiky (Rank Cases)

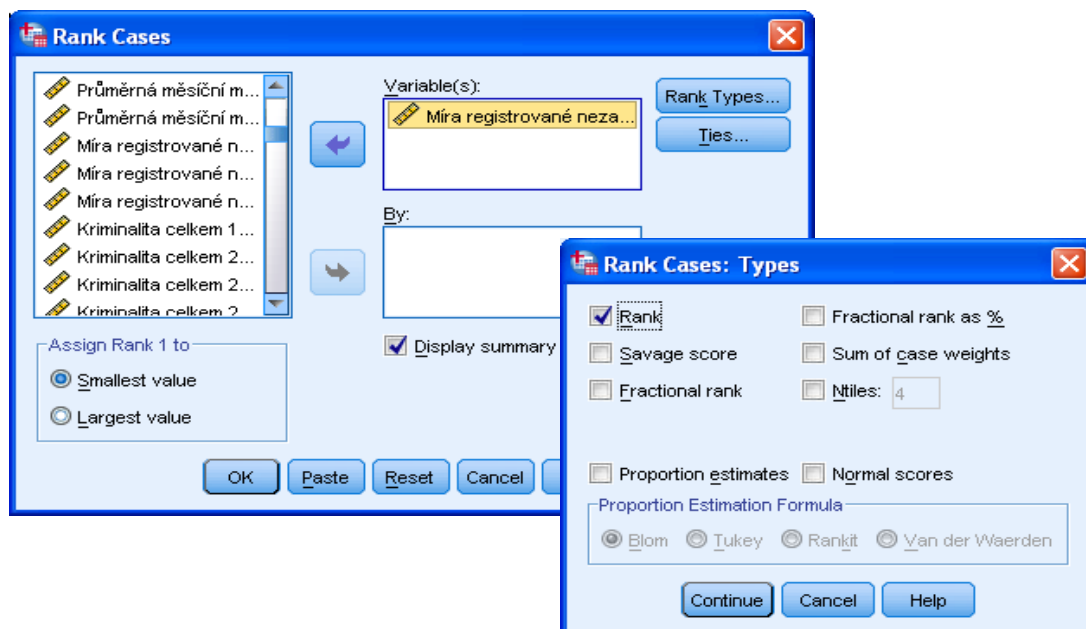
Procedura vytvoří proměnnou (proměnné), vyjadřující pořadí hodnot podle hrubých či normálních skóru dané numerické proměnné s přiřazením hodnoty 1 případu s nejvyšší nebo nejnižší hodnotou, případně s využitím procentuální škály podle zvoleného modelu poměrného odhadu.

Nejprve vybereme proměnnou (proměnné), ke které se bude pořadí vztahovat a převedeme ji šipkou do pole **Variable(s)**. Dále můžeme zvolit proměnnou, podle které se bude pořadí členit do podskupin, a převést ji do pole **By**.

V části **Assign rank 1 to** určujeme, zda má být hodnota 1 přiřazena případu s nejnižší (**Smallest value**) nebo nejvyšší hodnotou (**Largest value**).

Tlačítko **Rank Types** slouží pro výběr typu pořadí hodnot: pořadí (**Rank**), Savageův skór (**Savage score**), pořadí ve tvaru zlomku (**Fractional rank**), pořadí ve tvaru zlomku v % (**Fractional rank as %**), suma vážených případů (**Sum of case weights**), kategorizace podle kvantilů (**Ntiles**). Rovněž lze požadovat pořadové statistiky založené na kumulativní distribuční funkci (odhady proporcí (**Proportion estimates**) a normální skóry (**Normal scores**)).

Tlačítkem **Ties** určíme, jak přiřadit pořadí případům se stejnou hodnotou.



Výpočet pořadí hodnot

### Výpočet pořadových statistik

- Transform
  - Rank Cases

## 6.11 Průvodce pro práci s časem (Date and Time Wizard)

Procedura **Date and Time Wizard** je určena ke snadné práci s daty a časem. Po spuštění procedury se objeví průvodce, který naviguje k provedení požadovaného úkolu. Průvodce datem a časem nabízí tyto možnosti:

- **Learn how dates and times are represented in SPSS Statistics**

Stručný popis, jak IBM SPSS Statistics pracuje s daty (např. nelze vytvořit datum předcházející 14. říjnu 1582).

- **Create a date/time variable from a string containing a date or time**

Umožňuje vytvořit datum (čas) z textového řetězce, který musí obsahovat údaje interpretovatelné jako datum (čas), např. 12.1.1995. (Tato nabídka je aktivní pouze v případě, že v datech máme textovou proměnnou.)

- **Create a date/time variable from variables holding parts of date or time**

Vytvoří datum (čas) z více proměnných, které obsahují příslušné části data nebo času (roky, měsíce, dny, hodiny atd.)

- **Calculate with dates and times**

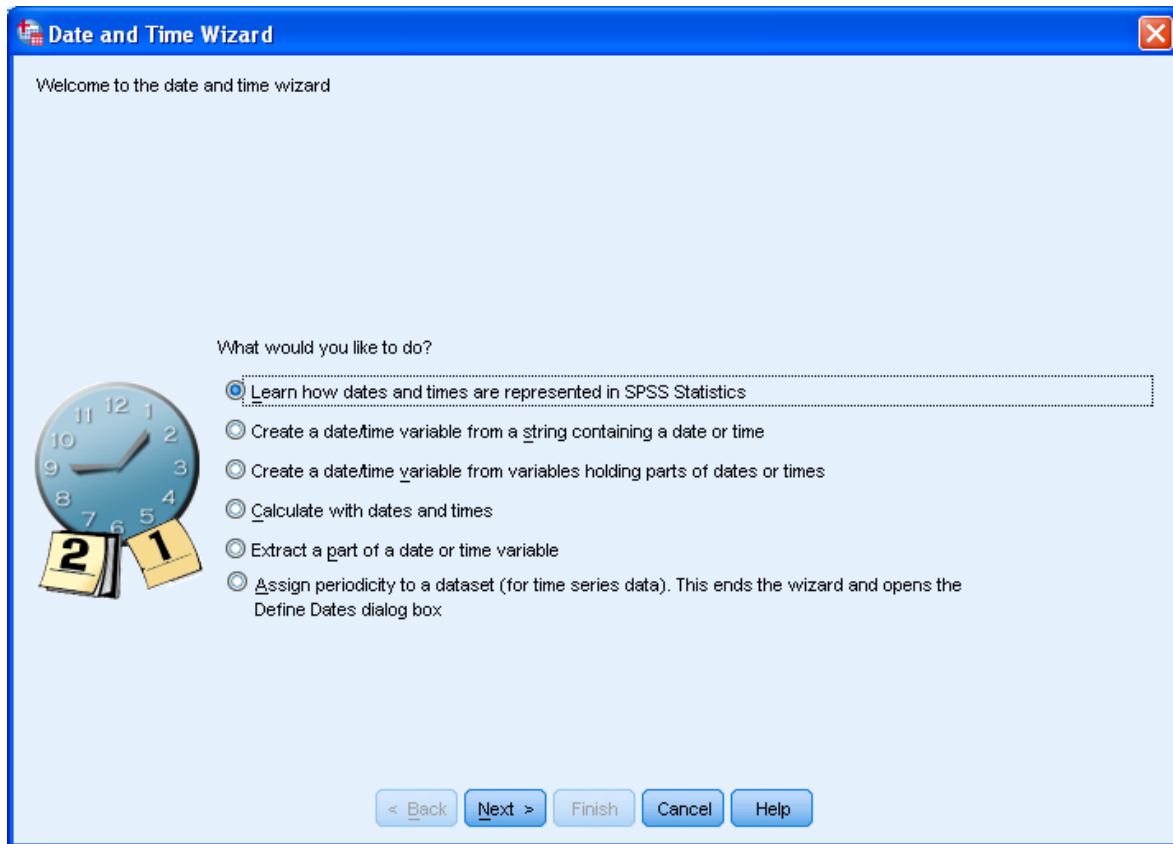
Umožní provést jednoduché výpočty s datem nebo časem: k datu (času) přičíst nebo odečíst konstantu či proměnnou, nebo spočítat rozdíl mezi dvěma daty (časy).

- **Extract a part of a date or time variable**

Umožňuje získat z data (času) libovolnou část, např. číslo měsíce. (Tato nabídka je aktivní pouze v případě, že v datech máme proměnnou typu „Date“.)

- **Assign periodicity to a dataset**

Vytvoří periodickou proměnnou zvolené délky. Tuto volbu je vhodné použít pouze v případě, že data představují časovou řadu.



*Průvodce pro práci s datem a časem*

### *Průvodce pro práci s časem*

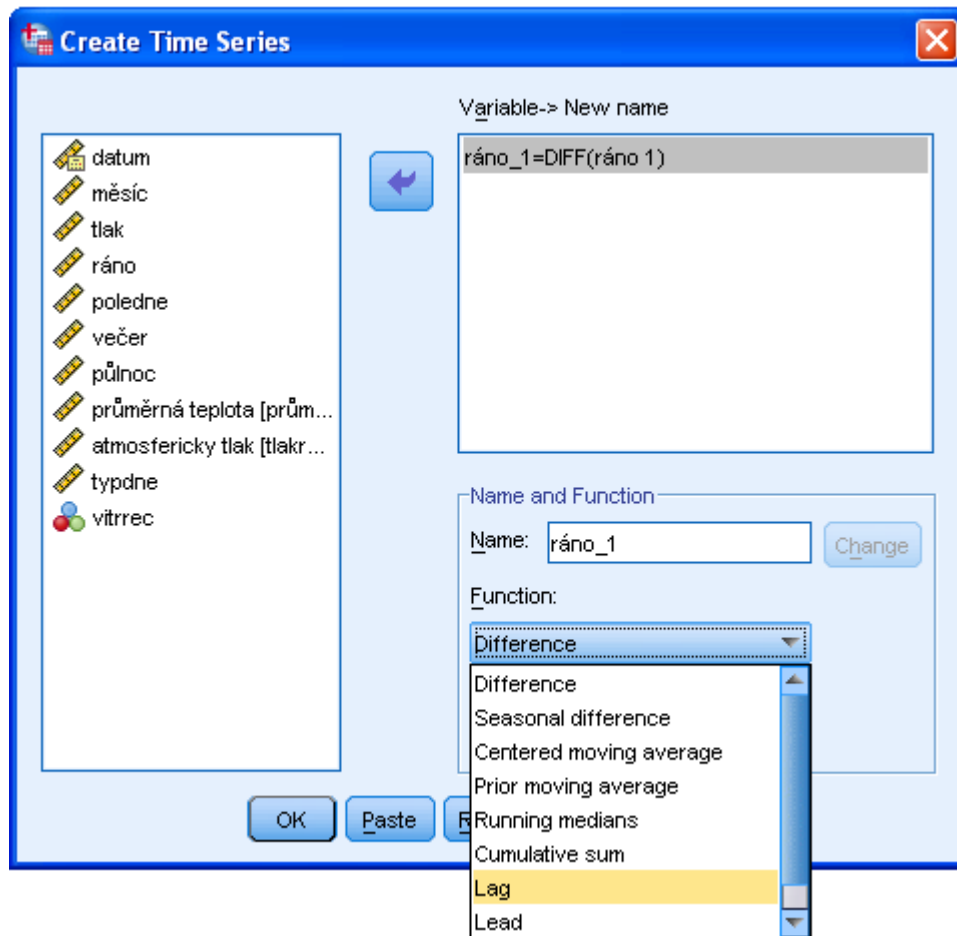
- *Transform*
  - *Date and Time Wizard*

## 6.12 Vytváření časových řad (Create Time Series)

Procedura vytvoří nové proměnné založené na existujících numerických proměnných s charakteristikou časových řad. Transformované hodnoty vzniknou aplikací zvolené metody na vybranou řadu (např. metoda vyhlazení časové řady pomocí klouzavých průměrů, rozdíl po sobě následujících pozorování, kumulativní součet apod.).

Nejprve zadáme proměnnou (proměnné) pro odvození nové časové řady a převedeme ji do pole **Variable** → **New name**. Rozbalovací seznam **Function** nabízí výběr z transformačních funkcí, které mohou být aplikovány na původní proměnnou. Novou proměnnou pojmenujeme v okénku **Name**. Zadání potvrdíme tlačítkem **Change**.

Pro některé funkce je dále nutné stanovit krok (**Order**) nebo rozpětí (**Span**).



Vytváření časových řad

### Vytváření časových řad

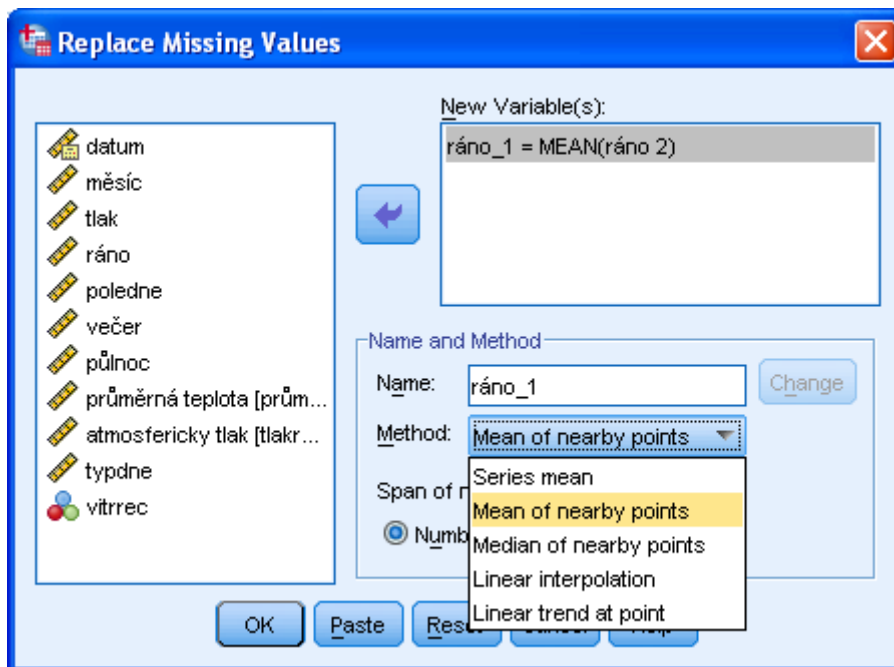
- Transform
  - Create Time Series

## 6.13 Nahrazení chybějících hodnot (Replace Missing Values)

Procedura vytvoří novou proměnnou (proměnné) tak, že u vybrané proměnné nahradí chybějící údaje zvolenou metodou (např. průměrem souboru, funkcemi okolních bodů v časové řadě, lineární interpolací apod.).

Do pole **New Variable(s)** přeneseme proměnné, u kterých chceme nahradit chybějící hodnoty. V poli **Name** zadáme název nové proměnné a dále zvolíme v rozbalovacím seznamu **Method** vhodnou metodu. Výběr potvrdíme tlačítkem **Change**.

Pro metody vycházející z funkcí okolních bodů je navíc nutné zadat rozpětí okolních bodů (**Span of nearby points**).



Nahrazení vynechaných hodnot

### Nahrazení chybějících hodnot

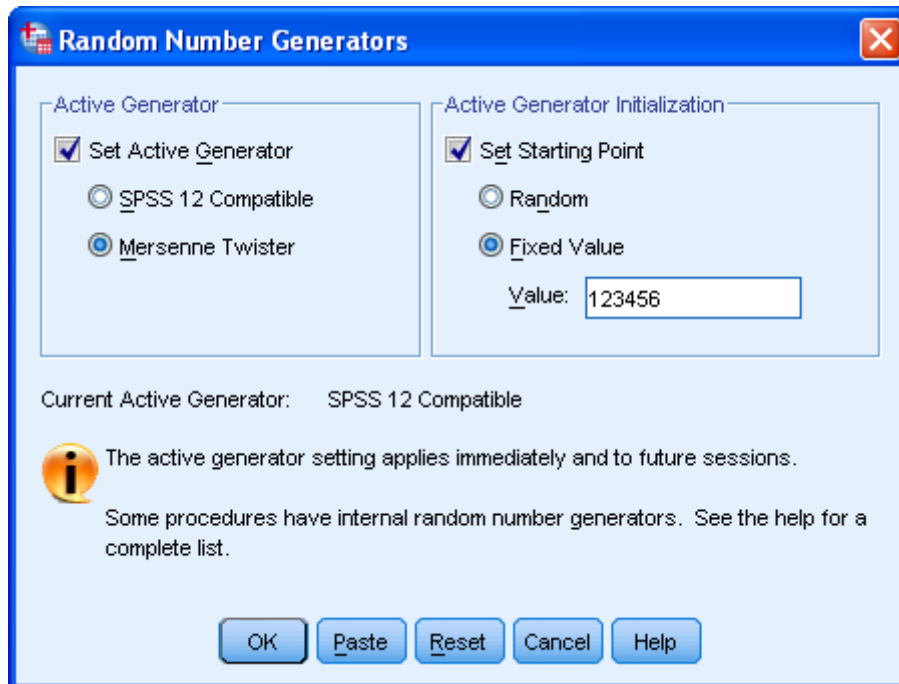
- Transform
  - Replace Missing Values

## 6.14 Nastavení generátoru náhodných čísel (Random Number Generators)

Tuto volbu můžeme využít, jestliže pracujeme s náhodným procesem (například náhodně vybíráme případy). V dialogovém okně volíme generátor náhodných čísel a případně také počáteční sekvenci čísel, což umožní kdykoliv později dojít při opakování tohoto náhodného procesu se stejným nastavením ke stejným výsledkům.

V části **Active Generator** určíme požadovaný generátor náhodných čísel. Jestliže potřebujeme z nějakého důvodu pracovat s generátorem, který byl v IBM SPSS Statistics k dispozici pro verze 12 a nižší, zvolíme **SPSS 12 Compatible**. **Mersenne Twister** představuje nový zdokonalený generátor.

V části **Active Generator Initialization** můžeme zadat počáteční hodnotu pro pozdější zopakování procesu (**Fixed Value**) nebo nechat začátek náhodný (**Random**).



*Nastavení generátoru náhodných čísel*

### ***Nastavení generátoru náhodných čísel***

- *Transform*
  - *Random Number Generators*

## **6.15 Provedení čekajících transformací (Run Pending Transforms)**

Tento příkaz je k dispozici pouze v případě, že je v menu **Edit** v nabídce **Options** na záložce **Data** zaškrtnuto **Calculate values before used**. Znamená to, že datové transformace nebudou provedeny ihned, ale pozdrží se do okamžiku, kdy zadáme z nabídky **Transform** příkaz **Run Pending Transforms**. Chceme-li přepsat data v okamžiku, kdy čekají nějaké transformace na provedení, program nás upozorní, že změny dat není možné uskutečnit a nabídne nám možnost realizovat nevyřízené transformace nyní.

Při obvyklém nastavení je v menu **Edit** v nabídce **Options** označena možnost **Calculate values immediately** a všechny transformace se potom provádějí ihned.

### ***Provést čekající transformace***

- *Transform*
  - *Run Pending Transforms*