

## 5. TRANSFORMACE PROMĚNNÝCH

### 5.1 Mazání proměnných a respondentů

Pro práci s jednotlivými případy a jednotlivými proměnnými má Stata dva komplementární příkazy: KEEP a DROP. Jejich výsledek je stejný, záleží proto jen na uživateli, který z nich v konkrétním případě použije.

**DROP** – smaže proměnnou nebo proměnné zadané za příkazem DROP.

```
drop year
drop year v1 v3
```

**KEEP** – ponechá proměnnou nebo proměnné zadané za příkazem KEEP, ostatní smaže.

```
keep year v1 v3
```

**DROP** – smaže případy splňující zadanou podmínku.

```
drop if year >= 2000
```

**KEEP** – ponechá případy splňující zadanou podmínku, ostatní smaže.

```
keep if age > 43
```

### 5.2 Změna proměnných a jejich obsahu

**RENAME** – změní jméno proměnné. Na prvním místě se za příkazem RENAME uvádí aktuální jméno, na druhém místě se uvádí nové jméno. Například následující příklad znamená „přejmenuj proměnnou, která se aktuálně jmenuje age; nové jméno nechť je vek“.

```
rename age vek
```

**REPLACE** – změní obsah proměnné (výpočtem). Příkaz se používá ve tvaru REPLACE *nazev\_promenne* = *vypocet*. Příkaz REPLACE je možno kombinovat s podmínkami IF.

```
replace vek = 1997 - v292
replace vek = 60 if vek > 60
```

**RECODE** – změní obsah proměnné (rekódováním). Po názvu proměnné nebo proměnných se uvádí přepisovací pravidla, která stanoví, která hodnota proměnné se přepíše na kterou novou hodnotu. Zajímavými parametry jsou COPYREST, který zkopíruje zbývající hodnoty, pro něž jsme nestanovili přepisovací pravidla, a GENERATE, který rekóduje výsledky do nové proměnné (viz dále). Pravidla se zadávají jednoduše pomocí znaménka rovná se (1 = 2) (2 = 3) (3 = 1). Rozsah „od do“ se zadává pomocí lomítka (1/3 znamená „od jedné do tří včetně“).

```
recode vek 0/15=1 16/65=2 66/110=3
recode vek 0 1 2 3 4 5 6 7 8 9 10 = 1, copyrest
recode vek 0/15=1 if year > 2010
recode v1 v2 v3 (1=5) (2=4) (3=3) (4=1) (5=1)
```

### 5.3 Vytváření nových proměnných

**GENERATE** – vytvoří novou proměnnou na základě výpočtu ze staré proměnné. Lze použít i jako parametr u příkazu recode. Před název nové proměnné můžete zadat její datový typ (byte, int, long, float, double).

```
generate vek2 = vek * vek
generate int vek2 = vek * vek
generate vek2 = vek * vek if rok == 2016
recode vek 0/15=1 16/34=2 35/100=3, generate (vek2)
```

**EGEN** – rozšířené vytváření nových proměnných. Tento příkaz je schopen vygenerovat novou proměnnou s využitím většího množství funkcí. Mezi zajímavé možnosti patří například funkce MIN, MAX, MEAN a MEDIAN, které generují novou proměnnou, jejíž všechny případy nabývají hodnot minima, maxima, průměru či mediánu jiné proměnné. Další možnosti skýtají funkce DIFF (vrací jedničku, pokud se hodnoty dvou proměnných v daném případě liší) nebo RANK (vrací pořadí konkrétní hodnoty v případě, kdy proměnnou seřadíte).

```
egen nej_vel = max (vek)
egen prum_vek = mean (vek)
egen rozdil = diff (v1 v2)
egen poradi_veku = rank (vek)
```