

4. INFORMACE O PROMĚNNÝCH

DESCRIBE – vypíše přehled všech proměnných nacházejících se v otevřeném datovém souboru včetně informací o jejich typu (byte, int, long – celá čísla, double, float – desetinná čísla, str – textový řetězec), popisu proměnné (*variable label*) a popisu jednotlivých hodnot (*value label*). Pokud vás zajímají informace o jedné konkrétní proměnné, napište její název za příkaz describe.

```
describe year
```

4.1 Datové typy

Proměnné ve Statě mohou mít jeden z následujících datových typů. Datový typ určuje, jakých hodnot může proměnná nabývat.

byte: celé číslo v rozsahu -127 až +100

int: celé číslo v rozsahu -32 767 až +33 740

long: celé číslo v rozsahu -2 147 483 647 až 2 147 483 620

double: desetinné číslo v rozsahu $-1,70141173319 \cdot 10^{38}$ až $1,70141173319 \cdot 10^{38}$

float: desetinné číslo v rozsahu $-8,9884656743 \cdot 10^{307}$ až $8,9884656743 \cdot 10^{307}$

str: textový řetězec, číslo za str udává počet znaků (např. str8 znamená osmiznakový textový řetězec)

Tip: pokud si nejste dopředu jistí, jaký typ proměnné použít, zvolte nejvyšší možný (long pro celá čísla, float pro desetinná čísla). Před uložením datového souboru pak použijte příkaz COMPRESS. Stata provede analýzu obsahu proměnných a sama zvolí nejvhodnější a nejúspornější datový formát.

4.2 Názvy a popisky

Stata rozlišuje několik názvů (name) a popisků (label) s následujícím významem:

Variable name: název proměnné ve formátu, ve kterém ho musíme zadávat při používání všech příkazů Staty. Nesmí se lišit ani ve velikosti písmen, ani v mezerách. Musí být zkratka zadán přesně tak, jak ho Stata zná.

Variable label: popisek proměnné. Slouží pouze pro lepší orientaci uživatele, Stata s ním nijak nepracuje, jen ho zobrazuje při výpisu výsledků namísto jména proměnné.

Value label: popisek hodnot proměnné. Pro možnost využívat statistické procedury jsou odpovědi respondentů tzv. kódovány do číselné podoby (např. pohlaví muž = 1, žena = 2, vzdělání ZŠ = 1, SŠ = 2, VŠ = 3 apod.). Pro snazší orientaci uživatele je možno přiřadit jednotlivým číselným hodnotám popisky (např. hodnotě 1 přiřadíme popisek „muž“). Tyto popisky se nikam nezadávají, Stata s nimi nepracuje, jen je zobrazuje při výpisu výsledků.

Tip: value label se často opakují (např. odpovědi na otázky Likertovy škály „zcela nesouhlasím“, „částečně nesouhlasím“, „ani nesouhlasím, ani souhlasím“, „částečně souhlasím“, „zcela souhlasím“). Proto Stata pracuje s popisky hodnot proměnných tak, že vytvoří pojmenovaný soubor popisků (např. výše uvedená škála by se mohla jmenovat LIKERTOVA_SKALA_POPISKY). Tento soubor popisků pak můžeme přiřadit několika proměnným.

LABEL – příkaz pro práci s popisky proměnných a hodnot proměnných. Používá se pro vytvoření popisku proměnné (label variable), definici souborů s popisky hodnot (label define) a přiřazení souboru s popisky hodnot konkrétní proměnné (label values). Pro výpis informací o popiscích hodnot slouží label list.

```
label variable kat_nejvzds "Kategorie nejvyššího dosaženého vzdělání"
label define KAT_NEJVZDS 1 "bez maturity" 2 "s maturitou" 3 "VŠ"
label values kat_nejvzds KAT_NEJVZDS
```

INSPECT – vypíše podrobnější informace o zvolené proměnné, konkrétně popisek proměnné, přehled rozložení hodnot, počet pozorování, počet chybějících hodnot a jednoduchý histogram.

```
inspect v306
```

CODEBOOK – vypíše informace o proměnných. Bez uvedení jména proměnné vypíše informace o všech proměnných v datovém souboru, při uvedení konkrétní proměnné nebo proměnných omezí výpis na informace pouze o zvolených proměnných. Výpis je možno upravit parametrem COMPACT pro zobrazení vybraných informací ve zkráceném formátu.

```
codebook
```

codebook v308, compact