

## Import danych panelowych do pakietu Stata mgr Karol Kuhl

Zbiór danych panelowych to zbiór, w którym wartości zmiennych dotyczą tych samych jednostek (krajów, osób, itp.), w kilku okresach (latach, miesiącach, itp.). Przykładem może być PKB *per capita* w poszczególnych krajach UE, w kolejnych latach. Każda obserwacja w zbiorze panelowym indeksowana jest podwójnie: po jednostkach (indeks *i*) i po czasie (indeks *t*). Dlatego dla każdej zmiennej można (np. w programie Excel) utworzyć następującą tabelę:

$(PKB \text{ per capita})_{it}$	2000	2001	...	2005
Austria	y	y		y
Belgia	y	y		y
...				
W. Brytania	y	y		y

Analogiczną tabelę można sporządzić dla stopy bezrobocia:

$r_{it}$	2000	2001	...	2005
Austria	x	x		x
Belgia	x	x		x
...				
W. Brytania	x	x		x

Pierwszym krokiem do importu danych panelowych do pakietu Stata jest zestawienie obydwu tabel (jeszcze w programie Excel):

$(PKB \text{ per capita})_{it}$	2000	2001	...	2005	$r_{it}$	2000	2001	...	2005
Austria	y	y		y	Austria	x	x		x
Belgia	y	y		y	Belgia	x	x		x
...					...				
W. Brytania	y	y		y	W. Brytania	x	x		x

Drugim krokiem jest prawidłowe przezwanie zmiennych:

kraj	y2000	y2001	...	y2005	x2000	x2001	...	X2005
Austria	y	y		y	x	x		x
Belgia	y	y		y	x	x		x
...								
W. Brytania	y	y		y	x	x		x

Taką tabelę można wkleić do pakietu Stata (oczywiście pamiętając o zamianie przecinków, będących separatorami dziesiętnymi, na kropki). Ta postać danych nosi nazwę "szerokiej" (*wide*). Pakiet Stata posiada polecenie umożliwiające transformację zbioru danych z postaci "szerokiej" na postać "długą" (*long*), konieczną do estymacji modeli ekonometrycznych (kluczową sprawą jest umieszczenie numeru okresu w nazwie zmiennych). Służy do tego polecenie:

```
reshape long y x, i(kraj) j(rok)
```

W poleceniu tym y i x są nazwami zmiennych, które podlegają transformacji; opcja i ( ) wskazuje na zmienną opisującą jednostki, a opcja j ( ) nadaje nazwę zmiennej opisującej lata. W wyniku tej transformacji otrzymuje się zbiór w postaci długiej:

kraj	rok	y	x
Austria	2000	y	x
Austria	2001	y	x
Austria	...	...	...
Austria	2005	y	x
Belgia	2000	y	x
Belgia	2001	y	x
...	...	...	...
Belgia	2005	y	x
...	...	...	...
W. Brytania	2000	y	x
W. Brytania	2001	y	x
...	...	...	...
W. Brytania	2005	y	x

Żeby na takim zbiorze móc estymować modele panelowe, należy wskazać zmienną numerującą jednostki oraz zmienną numerującą okresy (jest to odpowiednik polecenia `tsset` z analizy szeregów czasowych). Służą do tego polecenia:

```
iis kraj
tis rok
```

Dopiero teraz można szacować model postaci:

$$y_{it} = \alpha + \beta x_{it} + \gamma_i + \varepsilon_{it}$$

z tzw. efektami stałymi (`xtreg y x, fe`) lub losowymi (`xtreg y x, re`).