Ako naimportovať údaje do Gretlu

Gretl je voľne šíriteľný open-source ekonometrický softvér, ktorý budeme (hlavne z výukových dôvodov) používať ako hlavný nástroj na cvičeniach (popri iných profesionálnych nástrojoch). Okrem zjednodušenia, zrýchlenia práce a konečných výstupov ku ktorým sa chceme dopracovať poskytuje aj priebežné výstupy ktoré by sme museli v exceli pracne počítať a profesionálne štatistické softvéry ich už neuvádzajú. Samotný názov je akronym pre GNU Regression, Econometristic and Time-series Library.

Niektoré funkcie a ovládacie prvky programu sa môžu líšiť, vzhľadom na používanú verziu programu. V tomto prípade je uvádzaná verzia programu 1.9.5.cvs v anglickom jazyku, pričom je možné stiahnuť si program aj v iných jazykových mutáciách (dostupný aj v češtine).

Po otvorení programu sa zobrazí úvodné okno (Obrázok 1). Hlavné okno programu pozostáva z menu v hornej časti pomocou ktorého môžeme pracovať s údajmi, ďalej s hlavného okna v ktorom sa zobrazujú premenné s ktorými práve pracujeme (číslo premennej, názov a popis) a v ľavej dolnej časti okna sú ikony zrýchlujúce prácu a uľahčujúce prístup priamo k niektorým funkciám programu. Podobne ako v niektorých iných programoch, aj v Gretli je možné pracovať tiež pomocou konzoly a programovania, ale pre jednoduchosť budeme uvádzať len spôsob práce pomocou menu. V hlavnom menu na hornej lište sú možnosti File pre prácu so súbormi (otváranie, ukladanie, import, export súborov), Tools – táto možnosť poskytuje niektoré štatistické nástroje, štatistické tabuľky či testy, Data –možnosti práce s naimportovanými údajmi, View – možnosti grafického zobrazenia, či rôznych charakteristík, Add – možnosti pridávania premenných, či modifikácie dát, Sample-vytváranie a práca s výberovými súbormi, Variable – možnosti práce s jednotlivými premennými zvlášť, Model-ponuka ktorú budeme používať pri tvorbe ekonometrických modelov, Help – voľba poskytujúca manuál a základné informácie o programe.



Obrázok 1: Základné okno programu GRETL

Prvým krokom je vždy import údajov s ktorými chceme pracovať. Pokiaľ nemáme v programe naimportované údaje, možnosti práce v s údajmi v hlavnom menu zostávajú nevysvietené. Ak chceme do Gretlu naimportovať údaje, zvolíme v hlavnom menu možnosť file->open data-> na tomto mieste si môžeme vybrať či chcem otvoriť niektorý z pomocných súborov (sample file) ktoré program obsahuje, naimportovať nové údaje (Import), niektorý z uložených datasetov (User file) alebo otvoriť niektorý z posledný importovaných súborov (Obrázok 2).

Append data	E Sample file	
Save data as Export data Send To New data set Ctrl+N Clear data set Working directory Script files Session files Databases Eunction files	Import	text/CSV Octave Gnumeric Excel Open Document Eviews Stata SP55 SAS (xport) JMulTi

Obrázok 2: Otváranie a import údajov

ne I	ools Data J	iew Roo Saubie Tauabie Model Heib	
lo dataf	ile loaded		
D#	Variable name	Decorintive label	
_	greti: da	ta files	_ 0
	200		C
			1
	Greti Green	ne Gujarati Ramanathan Wooldridge	
	-		1
	File	Summary	
	Table_I.1	Personal consumption and GDP, US annual	
	Table_1.1	Egg production and prices, 1990 and 1991, by state	
	Table_1.2	CPI in seven industrialized countries, 1973-1997	
	Table_1.3	Exchange rates for seven countries, 1977-1988	
	Table_1.4	Seasonally adjusted M1 Money supply, guarterly 1959-1999	
	Table_1.5	Impact of Advertising Expenditure, by firm	
	Table_2.4	Family consumption and income	
	Table_2.6	Hourly wages by education (Goldberger)	
	Table_2.7	Labor force participation rates, men and women	
	Table_2.8	Food and total expenditure, 55 rural households, India	
	Table_3.2	Family consumption and income	
	Table_3.5	Ranks of students on midterm and final exams	
	Table_3.6	Productivity, U.S. business sector, 1959-1998	

Obrázok 3: Pomocné datasety dostupné v programe Gretl

V niektorých prípadoch budeme na cvičeniach využívať datasety v adresári Sample file. Do programu Gretl je možné doinštalovať datasety k príkladom z najznámejších svetových ekonometrických učebníc (Obrázok 3). V prípade zvolenia možnosti Sampe file sa otvorí okno v ktorom si na hornej lište môžeme vybrať buď databázu Gretlu, či súbory k učebniciam podľa autorov.

Pokiaľ sa rozhodneme importovať údaje, môžeme si v možnosti Import vybrať z niekoľkých typov dátových súborov ktoré je možné importovať do Gretlu. Sú tu všetky najčastejšie používané dátové formáty, pričom na cvičeniach budeme využívať prevažne možnosť importu údajov z hárkov programu MS Excel (Obrázok 2). Podobne ako pri importe údajov v iných štatistických softvérových nástrojoch, po tom ako si zvolíme možnosť importu údajov a typ dátového súboru ktorý chcem naimportovať, v ďalšom kroku nájdeme a vyberieme súbor podla jeho umiestnenia na pevnom disku (obrázok 4).

Places	Name	A Mo	dified
Recently Used	🐏 Opakovanie1-riešenie.xls	10:	09
 Jojo Desktop Vista (C:) Jednotka HD DVD-RO Data (E:) Jednotka DVD-RW (F:) Jednotka BD-ROM (G:) 	panei.xis	27.	

Obrázok 4:Cesta k importovanému súboru v strome dokumentov

Pokiaľ importujeme súbor MS Excel, mali by sme mať údaje pripravené tak aby boli v prvom riadku názvy premenných a údaje sa začínali prvým stĺpcom. Aby sme predišli možným problémom pri importe údajov je dobré nepoužívať v názve premenných interpunkciu a komplikovanejšie znaky. Po tom ako vyberieme umiestnenie importovaného súboru nám Gretl ponúkne výber hárku ktorý chceme zo zvoleného súboru naimportovať Na obrázku 5 vidíme, že importujeme súbor ktorý obsahuje 3 hárky pomenované Keynes, Spotreba piva a Produkcia.

Údaje by sa mali začínať v prvom stĺpci a prvom riadku, pokiaľ tomu tak nieje, v hornej časti okna nastavíme ktorým stĺpcom a riadkom sa začína import dátovej oblasti. V tomto prípade sme zvolili 1 stĺpec a 13 riadok na hárku Spotreba piva. Pokiaľ by sme si v tomto okne nič nezvolili, zostal by predvolený prvý hárok, na ktorom by sa začínal import údajov v prvom riadku a prvom stĺpci. Ak je všetko zvolené správne, po potvrdení prejdeme k ďalšiemu kroku v importe údajov, ak ste v tomto kroku zvolili nesprávnu možnosť Gretl ukáže chybové hlásenie.

💱 greti 📃	
E gret: spreadsheet import ariable Model Help	1
Nc Start import at:	
II column: 1 👘 row: 13 🖡	
(A)	
Sheet to import:	
Keynes	
Spotreba piva	
Cancel	
🖩 📝 🛅 🗰 fx 🔁 🔯 💒 β̂ 😫 🔚	

Obrázok 5: Výber hárku a umiestnenia dát

V ďalšom kroku je možné zvoliť charakter údajov, teda či sa jedná o prierezové údaje, údaje časových radov, prípadne panelové dáta. Pokiaľ v tomto kroku nenastavíme žiadny typ údajov, je možné že sa nebudú dať uskutočniť niektoré operácie špecifické pre istý druh dát. Ak v tomto kroku zvolíme typ údajov, bude sa dať ďalej s týmito údajmi ďalej pracovať už len ako so zvoleným typom dát. Pokiaľ by sme zvolili nesprávnu možnosť, typ údajov je možné zmeniť v hlavnom menu cez Data->Dataset structure.



Obrázok 6: Dialógové okno k typu údajov

Na obrázku 6 je dialógové okno v ktorom sa softvér pýta, či sa jedná o prierezové dáta alebo iný typ údajov. Ak by sme to potvrdili, softvér by pri ďalšej práci považoval importované údaje za prierezové, ak túto možnosť odmietneme, v ďalšom kroku môžeme zvoliť o aké údaje presne sa jedná. V tomto prípade sme importovali ročné údaje o spotrebe piva, teda ide o časové rady s ročnou periodicitou, takže v okne na obrázku 6 zvolíme možnosť No (nejedná sa o prierezové dáta). Po zvolení "No" dostaneme možnosť sami zvoliť o aké údaje sa jedná (Obrázok 7).

)#	Variable name	Descriptive label
0	const	auto-generated constant
2	DR	
4	PD	Data structure wizard
4	PR	Structure of dataset
5	I	Structure of dataset
6	v6	 Cross-sectional
		O Panel
		Fr
		Capcel

Obrázok 7: Voľba typu údajov

V tomto prípade ide o časové rady, takže zvolíme možnosť time series. Gretl nám následne ponúkne možnosť bližšie špecifikovať o aký časový rad sa jedná, teda o akú periodicitu údajov sa jedná a v akom období sa tento časový rad začína (Obrázo 8).

File .	Loois Data Mew	Data structure wizard	
ракоч	aniel-riesenie, xis	Time series frequency	
ID #	Variable name		
0	const		
11	0	O Quarterly	
2	PB	O Monthly	
4	PR		
5	I	O Weekly	
6	V6	O Daily (5 days)	
		🔿 Daily (6 days)	
		O Daily (7 days)	
		O Hourly	
		O Decennial	
		O Other	
		Cancel A Bark Conward	

Obrázok 8: Zvolenie frekvencie časového radu

Pokiaľ by sme zvolili prierezové údaje, softvér by len pre overenie vypísal koľko prierezových jednotiek v našom datasete našiel. Ak by sme zvolili Panelové dáta je niekoľko spôsobov ako ich naimportovať, najjednoduchším spôsobom je pripraviť si premenné ktoré rozlišujú jednotlivé prierezové jednotky a obdobia z ktorých pochádzajú pozorovania a pri importe tieto premenné zvoliť ako prierezovú a časovú premennú. V prípade úspešného importu údajov sa nám zobrazí hlavné okno, v ktorom sa objavia naimportované premenné. Po importe údajov sú v hlavnom menu dostupné už všetky možnosti (Obrázok 9).

🔵 gre	tl								_ 🗆 X
Eile	<u>T</u> ools	<u>D</u> ata	⊻iew	<u>A</u> dd	<u>S</u> ample	<u>V</u> ariable	Model	Help	
Opako	vanie1-	riešenie	.xls *						
ID #	Var	riable na	me	Descr	iptive labe	I			
0	сог	nst		auto-	generated	constant			
1	Q								
2	PB								
3	PL								
4	PR								
5	I								
	P-		fx 🕇	2 3	Αn	nual: Full r	ange 198	0 - 2009	

Obrázok 9: Gretl po úspešnom naimportovaní údajov

Pokiaľ by sme chceli naimportované údaje zmeniť či modifikovať existuje niekoľko spôsobov ako to uskutočniť. Pre modifikáciu a pridávanie ďalších premenných slúži možnosť Add v hlavnom menu. Prípadnú editáciu údajov je možné uskutočniť cez menu Data. Pokiaľ si chceme zobraziť údaje k niektorej premennej stačí dvojklik na príslušnú premennú. Ak chceme zobraziť či modifikovať údaje v celom datasete, zvolíme v ľavom dolnom okne štvrtú ikonu zľava ktorá poskytuje niektoré základné informácie a operácie s údajmi (Obrázok 10).



Obrázok 10: Základné informácie a operácie s údajmi

Po zvolení ikonky Dataset sa nám otvorí celý naimportovaný súbor (obrázok 11), ktorý môžeme ďalej modifikovať.

gretl: icon view	+	×	Q, 1983			
i			PB	PL	PR	
	1980	81,699997	1,78	6,95	1,11	
Data info	1981	56,900002	2,27	7,32	0,67	
161811	1982	64,099998	2,21	6,96	0,83	
15 (O)	1983	65,400002	2,15	7,18	0,75	
Summary	1984	64,099998	2,26	7,46	1,06	
3	1985	58,099998	2,49	7,47	1,1	
	1986	61,700001	2,52	7,88	1,09	
	1987	65,300003	2,46	7,88	1,18	
	1988	57,799999	2,54	7,97	0,88	
	1989	63,5	2,72	7,96	1,3	
	1990	65,900002	2,6	8,09	1,17	
	1991	48,299999	2,87	8,24	0,94	
	1992	55,599998	3	7,96	0,91	
	1993	47,900002	3,23	8,34	1,1	
	1994	57	3,11	8,1	1,5	

Obrázok 11: Zobrazenie všetkých naimportovaných údajov

Vypracoval: Jozef Palkovič