



# Corso di Analisi delle Serie Storiche a.a 2008 / 2009

## Laboratorio di Stata: lezione 2 – costruzione di datasets temporali, trasformazioni e analisi grafica



Carlo Drago  
[c.drago@mclink.it](mailto:c.drago@mclink.it)

5 dicembre 2008



# Riepilogo della scorsa lezione

---

- Introduzione a Stata
- Lo schermo iniziale
- Metodi di utilizzo di Stata
- La sintassi dei comandi
- Importazione dei dati
- Data editor e Data browser
- La guida in linea
- Creazione di serie storiche in Stata
- Salvataggio ed esportazione di dati
- Tipologie di files
- Creazione e modifica di variabili
- Statistiche descrittive



# Sommario

---

- La costruzione del dataset temporale
- Le trasformazioni delle variabili (1): variabili temporali, variabili indicatore.
- Le trasformazioni delle variabili (2): ritardi (lags), termini anticipativi (leads), differenze.
- Le trasformazioni delle variabili (3): medie mobili
- Le trasformazioni delle variabili (4): altre trasformazioni
- La descrizione dei dati e le statistiche descrittive
- L'analisi grafica dei dati



# Tipologie di datasets per serie storiche

A screenshot of a Stata spreadsheet showing a dataset with the following structure:

	A	B	C	D	E
1	ID	Mese	Anno	Valore	
2	1	1	2000	0.030674	
3	2	2	2000	0.059258	
4	3	3	2000	0.918936	
5	4	4	2000	0.180887	
6	5	5	2000	0.813101	
7					
8					
9					
10					

Two arrows point from the text below to the 'Mese' and 'Anno' columns.

Caso di più variabili temporali (una per mese ed una per l'anno). ID fornisce l'ordinamento delle osservazioni.

A screenshot of a Stata spreadsheet showing a dataset with the following structure:

	A	B	C	D	E
1	ID	Mese	Valore		
2	1	2000-1	0.705735		
3	2	2000-2	0.974566		
4	3	2000-3	0.299097		
5	4	2000-4	0.48682		
6	5	2000-5	0.736378		
7					

An arrow points from the text below to the 'Mese' column.

Caso di una variabile temporale (una sia per il mese che per l'anno).  
Notare che Stata riconosce anche 1-2000 come mese-anno (invece di 2000-1).  
ID fornisce l'ordinamento delle osservazioni



# La creazione di una chiave temporale: caso di più variabili temporali

---

- A meno che non si parta da un singolo variabile “anno” (in questo caso basterà utilizzare il comando **tsset**) la generazione di una variabile chiave avviene con il comando **gen** e passa attraverso 2 fasi: 1 trasformazione in edate delle variabili temporali (anno e mese ad esempio) 2 definizione della chiave temporale in Stata con **tsset**
- La procedura fa quindi uso specificamente di più colonne temporali che possono essere unite con l'apposita funzione in Stata, richiamata ancora una volta dal comando **gen** (ad esempio **ym** per anno e mese). Le funzioni sono nella slide successiva.
- A quel punto l'edate può essere trasformata con il comando **tsset** richiamando il corretto intervallo temporale.



# Creazione di chiavi temporali: funzioni (caso di più variabili temporali)

Funzione	Tipo di dati
<b><i>mdy(month,day,year)</i></b>	Giornalieri
<b><i>yw(year, week)</i></b>	Settimanali
<b><i>ym(year,month)</i></b>	Mensili
<b><i>yq(year,quarter)</i></b>	Trimestrali
<b><i>yh(year, half-year)</i></b>	Semestrali

Esempi:

**Gen data=mdy(mese,giorno,anno)** con mese, giorno, anno come tre variabili numeriche distinte

**Gen data=ym(anno,mese)**



# La creazione di una chiave temporale: caso di una variabile temporale

- La generazione di una variabile avviene con il comando `gen` e passa attraverso 2 fasi: 1 trasformazione in edate della stringa di testo (ad esempio 2000-1 per dati mensili) 2 definizione della chiave temporale in Stata con **tsset**
- La trasformazione permetterà di ottenere da una variabile stringa (la data di riferimento) un numero (tecnicamente l' "edate") e dunque trasformare la stringa iniziale che non sarebbe altrimenti trasformabile con il comando **tsset**. La trasformazione avviene mediante: **gen [nome]=funzione(variabile,opzioni)**
- Per i dati giornalieri partendo da una stringa (ad esempio 1/1/2003), potremo ottenere l'edate ad esempio con: **gen daily=date(sdate,"dmy")**
- A quel punto l'edate può essere trasformata con il comando **tsset** richiamando il corretto intervallo temporale.



# Creazione di chiavi temporali: funzioni (caso di una variabile temporale)

Funzione	Tipo di dati
<b><code>weekly(stringvar,"wy o yw")</code></b>	Settimanali
<b><code>monthly(stringvar,"my o ym")</code></b>	Mensili
<b><code>quarterly(stringvar,"qy o qy")</code></b>	Trimestrali
<b><code>halfyearly(stringvar,"hy o yh")</code></b>	Semestrali
<b><code>yearly(stringvar,"y")</code></b>	Annuali

Esempi:

**Gen data=monthly(mese,"ym")** con variabile costruita come 2000-1

**Gen data=monthly(mese,"my")** con variabile costruita come 1-2000



# La creazione di una chiave temporale: riepilogo

- La creazione di una chiave temporale in Stata avviene semplicemente facendo uso del comando **tsset** specificando l'intervallo temporale scelto. La trasformazione dall'originale tipo di variabili può avvenire mediante funzioni come **mdy()** laddove si parta da più variabili o da una funzione come **monthly()** laddove invece si parta da una singola variabile.
- Alla fine della procedura Stata comunicherà l'esistenza o meno di buchi (gaps) all'interno della serie.

Esempi:

**1) Gen data=mdy(mese,giorno,anno)**

**Tsset data, daily**

**2) Gen data=monthly(mese,"ym")**

**Tsset data, monthly**



# La creazione di ulteriori variabili

---

- Variabili anno, mese, settimana, giorno
- I ritardi (lags) e i termini anticipativi (leads)
- Trasformazioni in differenze
- Trasformazioni in media mobile
- Altre trasformazioni delle variabili (ad esempio il logaritmo facendo uso di **log()** )



# L'eliminazione delle variabili

- Il comando per eliminare una variabile è invece **drop**, laddove con **drop \_all** le elimineremo tutte.
- L'eliminazione di variabili è possibile anche mediante data editor, via il comando delete. E' sempre consigliabile avere un dataset salvato in diverso formato (ad esempio su un foglio di calcolo)

Delete

Data Editor

id[1] = 1

	id	anno	pi1	t	yhat	t2	yhat2	dum	yhat3	_est_model1	t3	yhat4	_est_model2
1	1	1870	53328.78	1	-257499	1	198809	0	146708.1	1	1	36050.18	1
2	2	1871	53937.56	2	-247454.1	4	187793.5	0	139910.6	1	8	39585.97	1
3	3	1872	53085.27	3	-237409.2	9	177104.6	0	133332.6	1	27	42930.51	1
4	4	1873	55155.11	4	-227364.3	16	166742.2	0	126974.2	1	64	46090.49	1
5	5	1874	55033.35	5	-217319.3	25	156706.3	0	120835.4	1	125	49072.64	1
6	6	1875	56616.17	6	-207274.4	36	146996.9	0	114916.1	1	216	51883.65	1



# La visualizzazione delle variabili trasformate

---

- Molto importante ricordare il comando **list** che permette di visualizzare ad una o ad una o a gruppi tutte le variabili (senza aprire né l'editor né il data browser).
- Facendo seguire a list la lista delle variabili cui siamo interessati possiamo rapidamente comparare i valori ottenuti. Con il comando **list \_all** visualizziamo tutte le variabili nel riquadro “variables”, senza aprire l'editor o il browser delle variabili stesse.



# Variabili anno, mese, settimana, giorno: la funzione **dofm()**

---

- Nel caso si avessero delle variabili già costruite (abbiamo effettuato la procedura di costruzione della chiave temporale usando due variabili) il problema non si pone, avremo già una variabile ad esempio che ci definisce il mese ed una l'anno. A quel punto si potrà direttamente costruire le variabili dummy (vedere più avanti).
- Nel caso avessimo una sola variabile data iniziale (ad esempio 1-2001) e avessimo in questa maniera costruito l'edate mediante la trasformazione in variabile (successivamente dichiarata chiave con **tsset**), potremo utilizzare una funzione ad hoc **dofm()** per la trasformazione in variabile mese, giorno o anno. Questa funzione permette di calcolare per esteso il numero dei giorni di ciascuna edate e dunque permette la creazione di variabili mensili, settimanali e così via.
- **Gen data2= dofm(data)**



# Creazione di variabili anno, mese, settimana, giorno

---

- La procedura, partendo da una sola variabile temporale, è dunque in due passi:
- **Gen data2= dofmm(data)**. Da una data iniziale si genera l'edate con i giorni numerati calcolati da Stata. Successivamente quindi possiamo creare la variabile mese, anno o giorno a seconda di quello che ci interessa facendo uso di tre funzioni di Stata **month()**, **year()** e **day()**. Stata in questo caso è in grado, avendo calcolato i corretti intervalli temporali di riconoscere il mese, l'anno ed il giorno di riferimento.
- **Gen mese=month(data2)** crea il mese di riferimento
- **Gen anno=year(data2)** crea l'anno di riferimento
- **Gen giorno=day(data2)** crea il giorno di riferimento



# La creazione di variabili indicatore (o dummy) temporali

- La creazione di variabili dummy può avvenire in vari modi, in particolare si segnala il comando **tabulate** (o **tab**).
- Il comando **tabulate (tab)** permette di ottenere le distribuzioni di frequenza per una data variabile. Ad esempio: **tab mese2, gen(d)**.  
permette di creare una variabile dummy per ciascuna osservazione.
- Ben più lunga (ma in alcuni casi estremamente utile) è la procedura di creare manualmente le variabili dummy facendo uso degli operatori di Stata. Si può quindi operare in questa maniera con **gen d=1 if mese==1** che crea una variabile d (dummy) uguale ad 1 laddove il mese sia 1, e visto che si otterranno dei valori mancanti (dove il mese sia diverso da 1) bisognerà far seguire un semplice **replace d=0 if d==.**



# I ritardi (lags)

Data  $y_t$  una serie storica possiamo avere i valori ritardati della serie stessa come  $y_{t-1}$ ,  $y_{t-2}$ ,  $y_{t-n}$ .

In Stata, data una variabile  $y$ , potremo ottenere i valori ritardati semplicemente utilizzando l'operatore ritardo o lag: **L.y** per  $y_{t-1}$ , **L2.y** invece per  $y_{t-2}$  e così via. In questo caso non è necessario creare una variabile.

Un modo alternativo è utilizzare i termini sottoscritti. avremo quindi **y[\_n-1]**, **y[\_n-2]** e così via, laddove però in questo caso dovremmo espressamente creare una variabile (utilizzando il comando **gen**).



# I termini anticipativi (leads)

Data  $y_t$  una serie storica possiamo avere i termini anticipativi della serie stessa come  $y_{t+1}$ ,  $y_{t+2}$ ,  $y_{t+n}$ .

In Stata avremo, data una variabile  $y$ , potremo ottenere i valori anticipati semplicemente utilizzando l'operatore forward: **F.y** per  $y_{t+1}$ , **F2.y** invece per  $y_{t+2}$  e così via. In questo caso non è necessario creare una variabile.

Un modo alternativo è utilizzare i termini sottoscritti. Avremo quindi  $y[_{n+1}]$ ,  $y[_{n+2}]$  e così via, laddove però in questo caso dovremmo espressamente creare una variabile (utilizzando il comando **gen**).



# Le trasformazioni in differenze

La trasformazione in differenze prime  $\Delta y_t = y_t - y_{t-1}$  si ottiene con l'operatore **D**. Le differenze seconde  $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$  si otterranno anch'esse con l'operatore **D2**.

Data ad esempio una variabile  $y$  in Stata, **D.y** permette di ottenere il termine della variabile  $y$  differenziato, laddove le differenze seconde, terze si ottengono con **D2.y**, **D3.y** e così via.



# Le variazioni di periodo

- Variazioni di periodo  $\Delta y_t = (y_t - y_{t-1}) / y_{t-1}$  possono essere facilmente ottenute a partire dall'operatore differenza, con **gen  $Vy=(D.y-D1.y)/D.y$**
- Variazioni relativamente a termini temporali precisi (ad esempio variazioni da gennaio a gennaio dell'anno precedente nel caso di dati mensili)  $\Delta y_t = (y_t - y_{t-4}) / y_{t-4}$  ad esempio, possono essere ottenute con l'operatore S. che si differenzia da D. in quanto mentre il secondo effettua la semplice differenza tra 2 osservazioni consecutive, S permette la differenza tra osservazioni non consecutive.
- Ad esempio **S4.y** è la differenza tra il valore attuale della variabile y e il valore ritardato di quattro periodi. Il comando è molto utile quando si voglia creare la variazione ad esempio ad 1 anno (**S12.y** nel caso mensile) e così via.



# Le trasformazioni in media mobile

- Le trasformazioni in media mobile:

$$y_t \rightarrow y_t^* = \sum_{i=-m_1}^{m_2} \mathcal{G}_i y_{t+i}$$

Pesi

$$m_1 \text{ e } m_2 \in \mathbb{N}; \mathcal{G}_i \in \mathbb{R}$$



# Le trasformazioni in media mobile

- Si possono ottenere delle medie mobili facendo uso sia di ritardi che di termini anticipatori. In Stata è possibile ottenere delle medie mobili sia via finestra di dialogo (il comando creato in questa maniera sarà **tssmooth**), sia manualmente creando delle variabili apposite.
- Ad esempio manualmente potremo avere: **gen y=(L.y+L1.y)/2** la media mobile a 2 periodi, a 3 periodi: **gen y=(L.y+L1.y+L2.y)/3** e così via. Si possono specificare manualmente anche i pesi, ad esempio ipotizzando: **gen y=0.6\*L.y+0.4\*L1.y** come media mobile pesata a 2 periodi, laddove 0.6 e 0.4 sono i pesi applicati.



# Le trasformazioni in media mobile

---

- All'interno della finestra di dialogo di **tssmooth** è possibile specificare i termini utilizzati nel calcolo della media mobile che possono comprendere termini ritardati, il valore corrente della serie storica e valori anticipativi. E' possibile anche specificare un peso da assegnare a ciascun termine.
- Come sempre è possibile effettuare modifiche sul comando generato dalla finestra di dialogo, opzione molto utile nel caso si vogliano creare più medie mobili e le si voglia confrontare.



# Le trasformazioni in media mobile

The screenshot shows the Stata 9.0 software interface. The 'Statistics' menu is open, and the path 'Statistics > Time series > Smoothers/univariate forecasters > Moving average filter' is highlighted. The 'Moving average filter' option is selected.

- Statistics
  - Summaries, tables, & tests
  - Linear models and related
  - Binary outcomes
  - Ordinal outcomes
  - Categorical outcomes
  - Count outcomes
  - Endogenous covariates
  - Selection models
  - Generalized linear models (GLM)
  - Nonparametric analysis
  - Time series**
    - Setup & utilities
    - ARIMA models
    - ARCH/GARCH
    - Prais-Winsten regression
    - Regression with Newey-West std. errors
    - Rolling windows estimation
    - Smoothers/univariate forecasters**
      - Single exponential smoothing
      - Double exponential smoothing
      - Holt-Winters nonseasonal smoothing
      - Holt-Winters seasonal smoothing
      - Nonlinear filter
      - Moving average filter**
    - Tests
    - Graphs
  - Multivariate time series
  - Longitudinal/Panel data
  - Survival analysis
  - Observational/Epi. analysis
  - Survey data analysis
  - Multivariate analysis
  - Resampling
  - Postestimation
  - Other



# Le trasformazioni in media mobile

Intercooled Stata 9.0 - [Results]

File Edit Prefs Data Graphics Statistics User Window Help

Review

drop dpil  
list pil  
list \_all  
edit  
edit  
edit  
help drop  
tssmooth ma my = pil, w  
list my  
edit  
drop my  
help tssmooth  
drop \_all  
edit  
tssset t  
tssmooth ma my = pil, w  
tssmooth ma vva = va, w  
list va vva  
tssmooth ma vva = va, w  
tssmooth ma vva2 = va, w  
tssmooth ma my = pil, w  
tssmooth ma vva2 = va, w  
use "H:\esercitazioni Stata\esercitazioni stata\_1\pil.dta"

Variables

d  
anno  
bil  
c  
what  
:2  
what2  
dum  
what3  
\_est\_model1  
:3

New variable è il nome della media mobile da creare, expression to smooth è invece l'espressione da creare in funzione ad esempio della variabile (basta indicare ad esempio la variabile di cui si vuole creare la media mobile).

tssmooth ma - Moving-average filter

Main f/in

Time Settings...

New variable: my = Expression to smooth: pil Create ...

New variable type: float

Replace existing variable

Number of lagged terms: 1

Current observation: Include (1) Exclude (0)

Number of lead terms: 1

Weight for current observations:

Weights for lagged terms: ?

Weights for lead terms: ?

OK Cancel Submit

Weight è invece relativo ai pesi da aggiungere eventualmente ai termini della media mobile.

Lagged terms sono I ritardi, mentre current observation è relativo alla possibilità di inserire il valore corrente, ove lead terms sono I valori anticipatori della variabile in questione da aggiungere o meno.

use "H:\esercitazioni Stata\esercitazioni stata\_1\pil.dta"



# Medie mobili di ordine diverso in Stata

- Si possono quindi creare medie mobili di ordine diverso in Stata, ad esempio medie mobili di ordine 3.

Ordine di una serie :

$$m_1 + m_2 + 1$$

Esempio ordine 3 :

$$m_1 + m_2 + 1 = 3$$

con  $m_1 = m_2 = 1$

$$y_t^* = \mathcal{G}_{-1}y_{t-1} + \mathcal{G}_0y_t + \mathcal{G}_1y_{t+1}$$

Stata permette di creare con **gen** e con **tssmooth** medie mobili differenti a seconda di come si specificano i termini ed i pesi. In questo caso manualmente avremo ad esempio una media mobile di ordine 3:

**Gen y=(F.y+y+L.y)/3**

o anche sempre di ordine 3:

**Gen y=(y+L.y+L2.y)/3**



# Media mobile semplice in Stata

Se i pesi sono tutti uguali e pari a :

$$\frac{1}{m_1 + m_2 + 1}$$

I pesi possono essere specificati via finestra di dialogo o manualmente. In questo caso si specificano i pesi manualmente. Essendo una media mobile semplice i pesi saranno tutti uguali ed avremo:

$$\text{Gen } y = (F.y + y + L.y) / 3$$



# Medie mobili centrate in Stata

- E' possibile in Stata creare anche medie mobili centrate

con  $m_1 = m_2 = m$

$$\frac{1}{2m+1} \sum_{i=-m}^m y_{t+i}$$

esempio ordine 3

$$m_1 + m_2 + 1 = 3$$

con  $m_1 + m_2 = 1$

$$y_t^* = \mathcal{G}_{-1} y_{t-1} + \mathcal{G}_0 y_t + \mathcal{G}_1 y_{t+1}$$

$$y_t^* = \frac{y_{t-1} + y_t + y_{t+1}}{3}$$

In questo caso, ad esempio, basta includere la variabile in questione all'interno della media mobile e fare uso degli operatori **F.** e **L.:**

$$\text{Gen } y = (\mathbf{F} \cdot y + y + \mathbf{L} \cdot y) / 3$$



# Le altre trasformazioni

---

- La trasformazione logaritmica della variabile mediante la funzione **log()** ottenibile ad esempio con il comando **gen ly=log(y)**



# Utilizzo dei range temporali nelle analisi

- E' possibile effettuare delle analisi "locali" tenendo conto soltanto di un certo insieme di osservazioni.
- Ad esempio in alcuni casi può essere utile considerare le osservazioni prima o dopo un determinato evento. E' possibile sempre fare uso di **if** (avendo specificato a priori un valore ID che indichi ciascuna osservazione), ma esistono due ulteriori modi specifici alle serie storiche in Stata.
- Per costruire range temporali: **If + tin()** oppure con alcune differenze **twwithin()**. Il comando viene quindi eseguito solo sul range temporale prescelto. Ad esempio **sum x if twwithin(2001-2,2001-7)** oppure **sum x if twwithin(2001-2,2001-7)**. Nel primo caso l'analisi sarà effettuata considerando anche i valori estremi nel secondo caso no.
- Per effettuare analisi su range temporali, specificata correttamente in Stata la data giornaliera, dovremo usare la condizione **If + w(), q(), y()** dopo il comando di riferimento. Ad esempio **sum y if (y2001)**
- Si rimanda comunque ai manuali ed alla guida in linea.



# La descrizione dei dati

---

- Descrizione dei dati (a livello informatico): **Describe**
- Descrizione dei dati in breve a livello statistico: **Codebook**
- Statistiche descrittive: **Sum**. Si possono ottenere delle statistiche descrittive più dettagliate con l'opzione **detail** (**Sum [variabile], detail**).
- Tutti questi comandi possono essere combinati con **by** e con **if** per ottenere delle statistiche descrittive per singolo anno, mese, e così via.
- I comandi di cui sopra seguiti da **\_all** permettono di generalizzare l'analisi a tutte le variabili contemporaneamente. Il medesimo risultato si può ottenere per gruppi di variabili facendo seguire alla prima variabile un trattino con la seconda e ultima variabile da analizzare (ad esempio **sum var1-var10** replica l'analisi da var1 a var10).



# I comandi codebook e describe

```
codebook pil

type: numeric (float)
range: [53085.27,1492840.6]          units: .01
unique values: 128                  missing.: 0/131
mean: 395421
std. dev: 443159
percentiles: 10%    25%    50%    75%    90%
              60390.6  81576  159621  589797  1.2e+06

describe pil

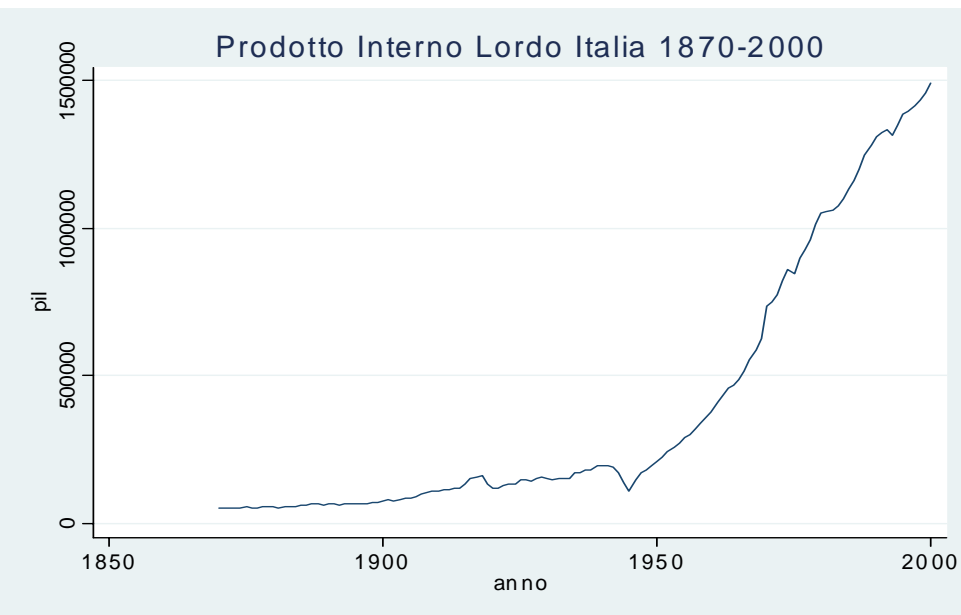
variable name  storage  display  value  variable label
              type   format   label
pil           float   %9.0g
```





# Visualizzazione grafica di serie storiche

- La visualizzazione della serie storica avviene tipicamente con il comando **Tsline [variabile 1]**
- Più serie storiche, possono essere visualizzate contemporaneamente con **Tsline [variabile 1, variabile 2, variabile 3]**





# Cenni di grafica in Stata

- I grafici in Stata possono essere fortemente personalizzati. Varie funzionalità permettono in un unico grafico di aggiungere più grafici visualizzati (comando **graph combine** trattato più avanti), permettono di aggiungere più tipologie di analisi statistica all'interno del medesimo grafico (ad esempio una retta di regressione su uno scatterplot come **twoway**). E' ulteriormente possibile sovrapporre più grafici nello stesso (ad esempio facendo uso nella finestra di dialogo di overlaid twoway graphs in Graphics).
- Stata permette altresì di creare degli oggetti visualizzabili sullo schermo come ad esempio delle linee di diverso colore all'interno di diagrammi temporali. Visualizzazione di linee sugli assi delle x o delle y: **tsline [variabile], xline() e yline()**
- A queste funzionalità si rimanda sui manuali e sull'help in linea.



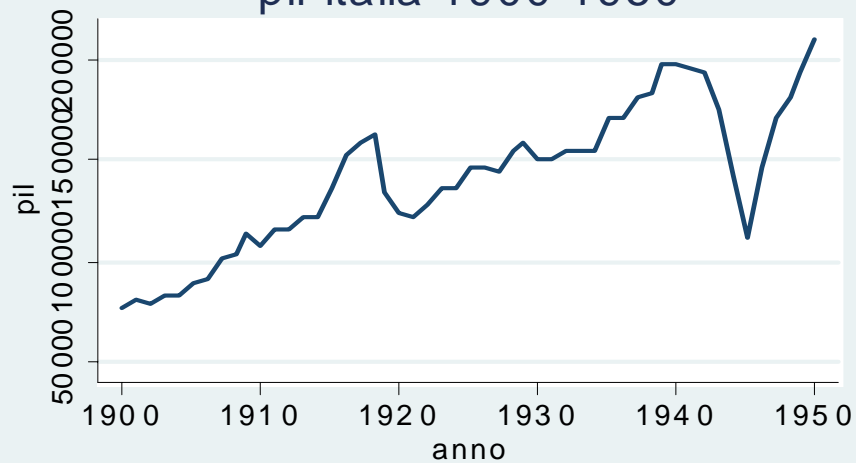
# Creazione di tavole di grafici

- E' possibile creare tavole di grafici semplicemente facendo uso del comando **Graph copy [nome grafico]** che permette di salvare uno o più grafici in memoria.
- Avendo salvato più grafici in memoria il comando **Graph combine [grafico1, grafico2...]** permette di creare tavole di grafici visualizzati contemporaneamente.
- In questo caso è assolutamente necessario controllare che la visualizzazione dei grafici sia corretta (ad esempio controllare la scala dei grafici e così via) e che gli stessi siano visibili. E' possibile modificare la struttura di un grafico dalla finestra di dialogo.

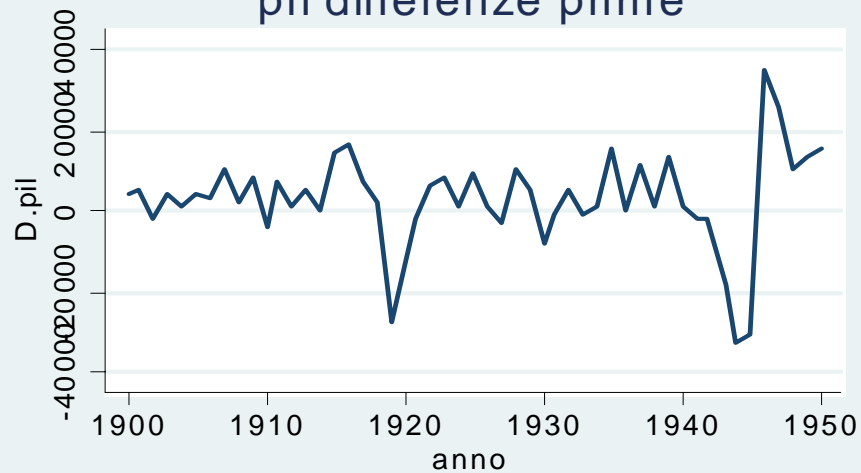


# Tavole di grafici

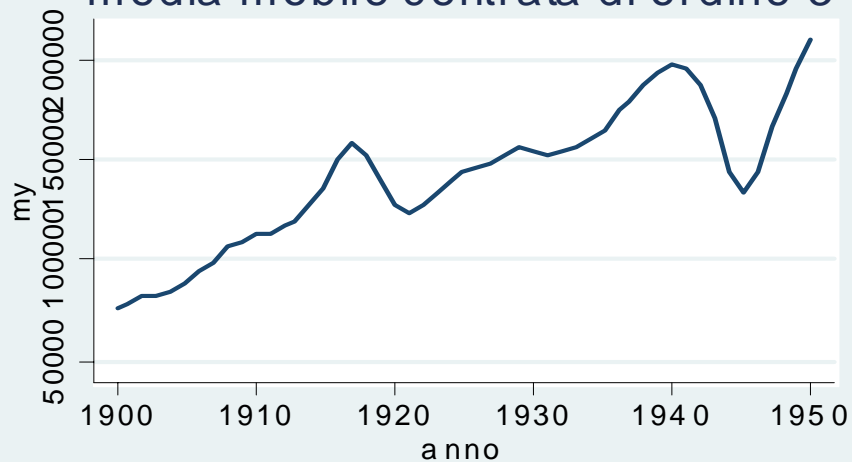
## pil italia 1900-1950



## pil differenze prime



## media mobile centrata di ordine 3





# Riepilogo

---

- La costruzione del dataset temporale
- Le trasformazioni delle variabili (1): (variabili temporali, variabili indicatore)
- Le trasformazioni delle variabili (2): ritardi (lags), termini anticipativi (leads), differenze.
- Le trasformazioni delle variabili (3): medie mobili
- Le trasformazioni delle variabili (4): altre trasformazioni
- La descrizione dei dati e le statistiche descrittive.
- L'analisi grafica dei dati



# Bibliografia

---

- Baum C.F. (2004) “Introduction to Stata”, Faculty Micro Resource Center Academic Technology Services, Boston College  
<http://fmwww.bc.edu/GStat/docs/StataIntro.pdf>
- Di Fonzo T., Lisi F. (2001) “Serie storiche economiche” Carocci editore
- Gartner M. (2000) “A primer in European Macroeconomics” Prentice Hall
- Gould W. (2003) “How do I create dummy variables?”  
<http://www.stata.com/support/faqs/data/dummy.html>
- Hardin J. (1996) “How do I create a lag variable?”  
<http://www.stata.com/support/faqs/data/lag.html>
- LSE Research Laboratory – IT Service (2004) “Introduction to Stata” [http://rlab.lse.ac.uk/it/it\\_docs/Introduction\\_to\\_stata.pdf](http://rlab.lse.ac.uk/it/it_docs/Introduction_to_stata.pdf)



# Bibliografia

---

- Milone G. (2005) “Laboratorio di Stata” Dipartimento di Matematica e Statistica, Università degli Studi di Napoli “Federico II”  
[http://www.docenti.unina.it/docenti/web/index.php?id\\_prof=331](http://www.docenti.unina.it/docenti/web/index.php?id_prof=331)
- Piccolo D. (2000) “Statistica” Il Mulino
- Scepi G. (2008) “Corso di Analisi delle Serie Storiche” Dipartimento di Matematica e Statistica, Università degli Studi di Napoli “Federico II”  
[http://www.docenti.unina.it/docenti/web/index.php?id\\_prof=331](http://www.docenti.unina.it/docenti/web/index.php?id_prof=331)
- Statacorp (2007) Stata 10 manuals
- Statacorp (2007) “Time Series Reference Manual”
- Statacorp website <http://www.stata.com/>



# Bibliografia

---

- Syracuse University Library (na) “Stata Tutorial”  
[http://library.syr.edu/information/mgi/nds/software/stata/intro/Reading\\_and\\_Documenting\\_Data.html](http://library.syr.edu/information/mgi/nds/software/stata/intro/Reading_and_Documenting_Data.html)
- Svend J. (2004) “Introduction to Stata”, Department of Epidemiology and Social Medicine, University of Aarhus  
<http://www.folkesundhed.au.dk/uddannelse/software/besked.pdf>
- UCLA Academic Technology Services (na) “Stata FAQ: How can I extract month and year component from a variable with %tm format?”  
[http://www.ats.ucla.edu/stat/stata/faq/time\\_funcs\\_ym2mandy.htm](http://www.ats.ucla.edu/stat/stata/faq/time_funcs_ym2mandy.htm)



# Per approfondire Stata e il calcolo statistico al computer (statistical computing)

---

- Baum C.F. (2002) “Why should you avoid using point-and-click method in statistical software packages?”  
<http://fmwww.bc.edu/GStat/docs/pointclick.html>
- Baum C.F (2005) “A little bit of Stata programming goes a long way” IDEAS Working Paper  
<http://ideas.repec.org/p/boc/usug05/16.html>
- Bookmark delicious sul calcolo statistico al computer ad UCLA  
<http://delicious.com/StatComp>
- Carolina Population Center: Stata tutorial  
<http://www.cpc.unc.edu/services/computer/presentations/stata/tutorial>
- Eszter’s Stata Goodies Page (Stata helpful resources)  
<http://www.eszter.com/stata.html>
- Portale su Stata ad UCLA <http://www.ats.ucla.edu/stat/stata/>



# Per approfondire Stata e il calcolo statistico al computer (statistical computing)

---

- Princeton University: Stata online training at DSS  
<http://www.princeton.edu/~otorres/Stata/>
- Ricerche consigliate su Google o altri motori di ricerca: Stata tutorial, Stata tips, Learning Stata, Introduction Stata, Stata lecture notes, Stata programs, Stata do files, Stata ado files
- Risorse per l'apprendimento di Stata presso Statacorp (raccolta di collegamenti) <http://www.stata.com/links/resources1.html>
- Risorse generali su Stata (programmi, esempi, datasets..) presso Statacorp <http://www.stata.com/links/>
- Statalist, gruppo di discussione su Stata  
<http://www.stata.com/statalist/>
- Stromberg, A.J. (2005) "Why write statistical software? The case of robust statistical methods", Journal of Statistical Software 10(5) <http://www.jstatsoft.org/v10/i05/paper>



# Riepilogo dei comandi di Stata (consultare l'help per maggiori dettagli)

- **Tsset** – definisce l'intervallo temporale per una serie storica
- **Gen** – genera la variabile
- **Drop** – elimina una variabile (specificando **\_all** le elimina tutte)
- **List** – visualizza una variabile (specificando **\_all** le visualizza tutte)
- **Tab** – distribuzioni di frequenza
- Operatori **L.** (ritardo) **F.** (anticipato) **D.** (differenza) **S.** (stagionale) da usare all'interno di una generazione di variabile con **gen**
- **Tssmooth** – trasformazione in medie mobili
- Funzione **log()** per trasformazioni in logaritmi da usare all'interno di una generazione di variabili con **gen**
- **Describe** – descrive il formato computazionale della variabile
- **Codebook** – statistiche descrittive di una variabile
- **Sum** – statistiche descrittive di una variabile. L'opzione **detail** permette di ottenere statistiche dettagliate.



# Riepilogo dei comandi di Stata (consultare l'help per maggiori dettagli)

---

- **By** ed **If** – esegue i comandi su partizioni del dataset specifiche
- **Graph copy [nome grafico]** – tiene in memoria il grafico salvato
- **Graph combine [nome grafico 1, nome grafico 2]** – crea una tavola di grafici