


## Lezione 23/3/09

Concatenamento di SAS data set  
Riorganizzazione di SAS data set

## Concatenamento SAS data set

- ◆ CONCATENAMENTO UNO A MOLTI  
(merge2.sas)
- ◆ FUSIONE UNO A UNO CON BY E IN=   
(merge3.sas)
- ◆ FUSIONE UNO A MOLTI CON BY E IN=  
(merge3.sas)
- ◆ FUSIONE MOLTI A MOLTI CON BY E IN=  
(merge3.sas)

lucidi SAS 113-119



## Opzione IN=

```
SET dati1 (IN=var) dati2;
```

La variabile **var** consente di individuare il data set da cui provengono i dati presenti nel PDV ad ogni iterazione del data step:

se l'osservazione proviene da **dati1**  
**var=1**, altrimenti **var=0**



## Ordinamento degli archivi

- ◆ Procedura SORT


```
PROC SORT [DATA=dataset] [OUT=dataset2];  
BY [DESCENDING] var1[[DESCENDING] var2];
```

## Esercizio 1: utilizzo di IN=

- ◆ Una ditta produttrice di articoli sportivi vuole inviare un depliant pubblicitario a tutti i clienti che non hanno effettuato ordini durante l'anno in corso.
- ◆ L'azienda dispone di due archivi di dati, uno contenente l'elenco dei clienti (CLIENTI.DAT) e l'altro l'elenco degli ordini (ORDINI.DAT).
- ◆ L'archivio clienti contiene le seguenti variabili: CODICE CLIENTE, NOME, INDIRIZZO
- ◆ L'archivio ordini contiene una osservazione per ogni ordine effettuato nell'anno in corso e, per ciascun ordine, le variabili: CODICE CLIENTE, IMPORTO ORDINE

INT: unire i due archivi utilizzando l'opzione IN= e selezionare i clienti che non hanno osservazioni nell'archivio ordini. (soluzione in [clienti.sas](#))

## Riorganizzazione SAS data set

- ◆ Variabili in osservazioni: [reshape1.sas](#)
- ◆ Valori di sintesi in osservazioni: [reshape2.sas](#)
- ◆ Valori di sintesi per gruppo in osservazioni: [reshape3.sas](#) 

## Riorganizzazione SAS data set

- ◆ Valori di sintesi per gruppo in osservazioni: [reshape3.sas](#)
  - sas-data set REGIONI (info per regione)
  - calcola media di alcune variabili per ripartizione con PROC MEANS+BY;
  - Unione archivio originale con valori di sintesi per ripartizione: [MERGE+BY](#)

[reshape3.sas](#)

## Fusione con variabili comuni

Se ci sono valori ripetuti rispetto alle variabili comuni



nel nuovo data set si considera l'osservazione ripetuta in tutti gli archivi

[esercizi in merge.pdf](#), [mergedoc.sas](#)

## Esercizio 2

- ◆ Considerare i due file VARESE\_ANAGR.CSV e FRAZIONI\_VA\_ANAGR.CSV.
- ◆ Importare i due file e si crei un archivio in cui si considerano i due archivi concatenati (tramite SET).
- ◆ Assegnare le etichette (LABEL) alle variabili "cod\_com" (Codice del comune) e a Cod\_fam (Codice della famiglia).
- ◆ Ordinare i dati in base al numero di componenti (ncomp).
- ◆ Stampare i risultati.

## Esercizio 3

- Importare i files in formato .txt dal nome "Squadre" e "Ruolo" con le seguenti variabili:
  - ◆ Squadre: Cognome Nome eta anni\_gioco cambio\_squadre
  - ◆ Ruolo: Cognome ruolo squadra\_attuale
- ◆ Unire i due file (fusione) tramite il comando MERGE.
- ◆ Costruire una nuova variabile chiamata "mobilita" con i valori seguenti:
  - 1 se le squadre cambiate sono <3
  - 2 se le squadre cambiate sono >=3
- ◆ Ordinare le osservazioni sulla base della variabile appena creata e stampare i risultati.

## Esercizio 4

- Partendo dai dati contenuti nei files COLLEZIONE\_INT.CSV e COLLEZIONE\_NAZ.CSV in formato CSV:
  - ◆ Si importino i due files e si crei un unico archivio tramite il comando SET.
  - ◆ Si stampino le osservazioni separandole per negozio di appartenenza.
  - ◆ Si importino i due files e si crei un unico archivio tramite il comando MERGE utilizzando la variabile "Negozio" come chiave di unione.
  - ◆ Si stampino le osservazioni del nuovo archivio.

## Ricodifica domande doppie

- Analisi delle risposte ai due quesiti seguenti:
  - ◆ *Indicare gli elettrodomestici che intende acquistare nei prossimi 12 mesi:*
    - Frigorifero
    - Lavatrice
    - Lavapiatti
  - ◆ *Indicare dove intende acquistare gli elettrodomestici:*
    - Ipermercato
    - Grande magazzino
    - Negozio specializzato