

## Esercitazione 4/5/2009

### Esercizi con PROC TABULATE

## Esercizio 1

- ◆ Si importi il file in formato csv dal nome VARESE\_ANAGR.CSV.
- ◆ Si costruiscano con PROC TABULATE le seguenti tabelle
  - somma del numero di componenti per TIPOFAM (righe) e TIPO\_NUCLEO (colonne)
  - numero medio di componenti per TIPOFAM (righe) e TIPO\_NUCLEO (colonne)
  - Distribuzione di frequenza congiunta per TIPOFAM (righe) e TIPO\_NUCLEO (colonne)

## Esercizio 2

Importare il file in formato csv COUNTRY.CSV

- ◆ In un unico un passo di data:
  - costruire la variabile RICCHEZZA con le seguenti modalità:
    - ◆ Se  $GDP < 2000$  allora ricchezza=bassa
    - ◆ Se  $GDP \geq 2000$  allora ricchezza=alta.
  - costruire la variabile DIMENSIONE che vale 1 se la popolazione al 92 è inferiore a 10 e 0 altrimenti.

Si stampino i primi 10 record del data set creato.

## Esercizio 2 segue

- Disegnare i box-plot di LIFEEXPM in base alle modalità di RICCHEZZA
- Utilizzando una procedura opportuna, si calcoli il valore medio condizionato di LIFEEXPM in base alle modalità di RICCHEZZA
- Costruire la tabella contenente il valore medio di LIFEEXPM e LIFEEXPF e le frequenze assolute per DIMENSIONE (variabile di riga) e RICCHEZZA (variabile di colonna) con indicazione dei totali di riga e di colonna.
- Per i soli paesi con DIMENSIONE=1, costruire la tabella contenente frequenze assolute (formato 6.0), percentuali (formato 6.2) sul totale e valore medio di GDP (formato 6.2) generale di RICCHEZZA, con indicazione del totale

## Esercizio 3

- ◆ Riprendere il data set creato nell'esercizio del 17 maggio 2007 relativo all'indagine sui mezzi di trasporto (soluzione in v:\carla\sas\ esercizio17\_5\_07.sas)
- ◆ Definire opportuni formati per le variabili da utilizzare nei successivi punti 2 e 3. In particolare, definire 4 classi di età in base ai quartili della distribuzione.

## Segue esercizio 3: utilizzo di PROC TABULATE

- ◆ costruire la tabella contenente il numero medio di corse, utilizzando come variabili di riga (incrociate) la professione e il sesso e come variabile di colonna le classi di età;
- ◆ costruire la tabella contenente il numero di intervistati e il numero medio di corse per sesso (variabile di riga) e utilizzo del mezzo pubblico (variabile di colonna), la distribuzione % degli intervistati per sesso e la distribuzione % delle corse per utilizzo del mezzo pubblico.