

Statistica computazionale
Esercizi istruzione IF

ESERCIZIO 1

Il file ASCII denominato V:\CARLA\SAS\COMPITI\SAS6_2_02.TXT contiene la valutazione di alcuni insegnamenti da parte degli studenti iscritti presso l'Ateneo YY per l'A.A. 2000/01. Le variabili che interessano e la posizione da esse occupata nel record sono riportate nella seguente Tabella 1.

Tab. 1 – Variabili archivio SAS6_2_02.txt

posizione	Contenuto
1-3	Codice insegnamento
5	Sesso (1=maschio, 2=femmina)
7	Classe di età (1-4)
9-11	Voto di maturità (anno 1999: 60-100) per i maturati dal 1999
13-14	Voto di maturità (anno precedente '99: 36-60) se maturati prima del '99
16-17	Numero di corsi frequentati
19	Chiarezza del docente (1=No, 2=più no che sì, 3=più sì che no, 4=sì)
20	Basi adeguate (1=No, 2=più no che sì, 3=più sì che no, 4=sì)
22	Interesse per la disciplina (1=No, 2=più no che sì, 3=più sì che no, 4=sì)
24	Soddisfazione globale (1=No, 2=più no che sì, 3=più sì che no, 4=sì)

Scrivere un "programma SAS" che inizi con una riga di commento contenente il nome del file che conterrà il programma stesso e che svolga le seguenti operazioni in un unico passo di DATA:

- 1 crei due data set SAS temporanei: il primo contenente i dati relativi ai maschi e l'altro i dati relativi alle femmine. Nell'archivio dei maschi non va inserita la variabile relativa al codice dell'insegnamento;
- 2 eliminare dall'archivio delle femmine gli studenti della classe di età 3;

in entrambi gli archivi

- 3 creare un'unica variabile per il voto di maturità, VOTOMAT, con voti espressi da 60 a 100 e quindi eliminare le due variabili di partenza;
- 4 creare una variabile dicotomica (0,1), che assuma valore 1 se lo studente è interessato alla disciplina (codici 3 e 4) e 0 altrimenti;
- 5 creare una variabile che sia pari al logaritmo del numero di corsi frequentati;
- 6 assegnare una descrizione agli archivi;
- 7 assegnare una descrizione alle variabili.

- Stampare il contenuto della DIA di entrambi gli archivi;
- Stampare le prime 10 osservazioni dell'archivio relativo alle femmine, escludendo dalla stampa le osservazioni relative alle studentesse con voto di maturità inferiore a 90;
- Stampare le osservazioni dalla seconda alla dodicesima dell'archivio maschi, stampando solo le variabili relative al numero di corsi frequentati e al giudizio sulla chiarezza del docente.

ESERCIZIO 2

Il file "MONDO95.csv" contiene i dati relativi ad alcuni indicatori per paese misurati nel 1995. In particolare le variabili contenute nell'archivio sono le seguenti

- Nationality
- Cod.
- Population in thousands
- Number of people / sq. kilometer
- People living in cities (%)
- Predominant religion
- Average female life expectancy
- Average male life expectancy
- People who read (%)
- Population increase (% per year)
- Infant mortality (deaths per 1000 live births)
- Gross domestic product / capita
- Region or economic group (codici: 1 Developed Country, 2 Est Europe, 3 Pacific/Asian, 4 Africa, 5 Middle Est, 6 Latin America)
- Daily calorie intake
- Aids cases
- Birth rate per 1000 people
- Death rate per 1000 people
- Number of aids cases / 100000 people
- Log (base 10) of GDP_CAP
- Log (base 10) of AIDS_RT
- Birth to death ratio
- Fertility: average number of kids
- Log (base 10) of Population
- Males who read (%)
- Females who read (%)
- Predominant climate (codici 1 Desert, 2 Arid/Desert, 3 Arid, 4 Tropical, 5 Mediterranean, 6 Marittime, 7 Temperate, 8 Artic/Temp, 9 Artic)

1. Importare il file.
2. Assegnare un titolo al programma.
3. Ricodificare (tramite l'istruzione IF) le due variabili REGIONE e CLIMA che nel file originale sono in formato numerico, attribuendo loro la stringa sopra definita.
4. Costruire una nuova variabile (da chiamare CONFRONTO) di confronto tra la speranza di vita per gli uomini e per le donne come rapporto tra VITAMAS/VITAFEM*100.
5. Assegnare una nuova variabile (SPERANZA) con i valori seguenti:
 - Maggiore speranza maschile se CONFRONTO>100
 - Maggiore speranza femminile se CONFRONTO<100
 - Uguaglianza se CONFRONTO=100
6. Tenere solo le variabili create e la variabile NAZIONE.
7. Ordinare i dati in base alla variabile SPERANZA.
8. Stampare ad OUTPUT i risultati

Esercizio 3

L'archivio "Prod.patate.csv" è in formato csv.

- Importare il file e salvare le istruzioni generate da 'import wizard' nel file che contiene il programma.
- Creare la variabile "raccolta" con le due modalità seguenti:
 - se Perc_racc<90, allora raccolta=bassa;
 - se Perc_racc>=90, allora raccolta=alta.
- Tramite le istruzioni IF...DO si crei una variabile che assegni 1 se raccolta=alta e altro altrimenti.
- si stampino il contenuto dell'archivio e quello della DIA.