

INTRODUZIONE AL SAS

Rapporto Interno C86-20

15 Dicembre 1985

Rolando Bianchi Bandinelli
Edoardo Bracci
Renato Ferrini
Domenico Laforenza

INTRODUZIONE AL SAS

(STATISTICAL ANALYSIS SYSTEM)

PISA, 19-21 Dicembre 1985

Rolando Bianchi Bandinelli
Edoardo Bracci
Renato Ferrini
Domenico Laforenza

GNUCE

Istituto del Consiglio Nazionale delle Ricerche

PROGRAMMA CORSO SAS BASE

- I1 - Introduzione alla programmazione
Da cosa si parte (DATI); processo di elaborazione (ALGORITMO); cosa si ottiene; come si ottengono i risultati.
- I2 - Presentazione del SAS
Facilita di uso; interattivo e batch; integrazione della statistica, della grafica ecc.
- C1 - Considerazioni generali sul SAS
Storia; prodotti aggiuntivi; bibliografia; stato dell'arte; SUGI
- C2 - Funzionamento del SAS
DATA; Input buffer; Files; INPUT; CARDS; PROC PRINT
- C3 - Data step - fondamentali
DATA; INPUT; Assegnazione; RETAIN; SUM; IF-THEN-ELSE; DO-END; DELETE; IF subsetting
DROP-KEEP; Functions; SET
- C4 - PROCEDURE step - fondamentali
Istruzioni: VARIABLES; BY; DROP; ecc.
Procedure: PRINT; CONTENTS; CHART; MEANS; FREQ; PLOT; FORMAT
- C5 - Esempi di programmi SAS
- C6 - Ulteriori istruzioni SAS
SET-MERGE-UPDATE; gestione date e tempi; LABEL; FORMAT-INFORMAT; gestione dei files non SAS; FILE-PUT-INFILE

II

INTRODUZIONE

ALLA

ELABORAZIONE ELETTRONICA

GENERALITA'

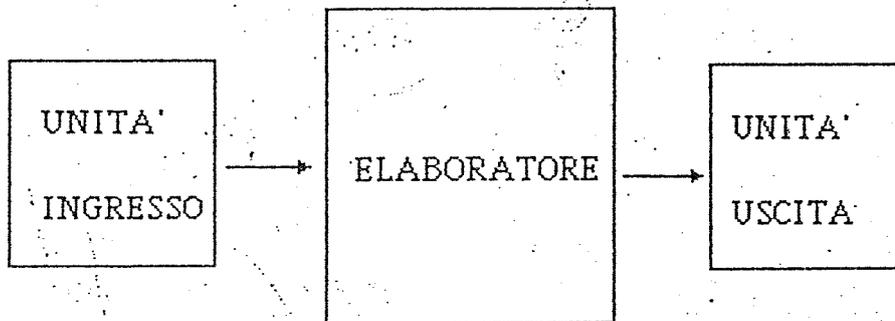
UN SISTEMA DI ELABORAZIONE SI COMPONE DI TRE GRUPPI
DI COMPONENTI:

- HARDWARE: INSIEME DEI DISPOSITIVI MECCANICI ED
ELETTRONICI CHE COSTITUISCONO UN
ELABORATORE (UNITA' CENTRALE DI
ELABORAZIONE, MEMORIA ECC.)

- SOFTWARE: INSIEME DEI PROGRAMMI CHE PERMETTONO
DI FAR FUNZIONARE UN ELABORATORE

- I/O: INSIEME DI DISPOSITIVI CHE COLLEGANO
UN ELABORATORE AL MONDO ESTERNO
(LETTORI DI SCHEDE, TERMINALI, UNITA'
NASTRO, UNITA' DISCO, STAMPANTI ECC.)

GENERALITA'



UNITA' INGRESSO : SCHEDE

TERMINALE

DISPOSITIVI MAGNETICI

UNITA' USCITA : STAMPANTE

TERMINALE

DISPOSITIVI MAGNETICI

ELABORATORE : ESEGUE UNA SERIE DI ORDINI

ELEMENTARI (ISTRUZIONI) CHE

COSTITUISCONO UN PROGRAMMA

PROGRAMMA

PROGRAMMA : INSIEME ORDINATO DI ISTRUZIONI FATTE
ESEGUIRE AD UN ELABORATORE FINO AL
TERMINE DI UN DETERMINATO LAVORO

LAVORI DIVERSI RICHIEDONO PROGRAMMI DIVERSI

I PROGRAMMI SONO SCRITTI IN LINGUAGGI CHE SONO
DIVENTATI SEMPRE PIU' VICINI AL LINGUAGGIO NATURALE
IN MODO DA POTER ESSERE ADOPERATI DA UN SEMPRE
MAGGIOR NUMERO DI UTILIZZATORI NON NECESSARIAMENTE
SPECIALISTI IN INFORMATICA

PER CONVERTIRE I PROGRAMMI NEL LINGUAGGIO
DELL'ELABORATORE (CERTAMENTE POCO UMANO !) OCCORRONO
DEGLI STRUMENTI IN GRADO DI COMPIERE QUESTA TRADUZIONE
QUESTI STRUMENTI SI CHIAMANO "COMPILATORI".

OLTRE ALLA TRASFORMAZIONE DEI PROGRAMMI, ESSI COMPIONO
ANCHE UNA ANALISI FORMALE SEGNALANDO EVENTUALI
ERRORI DI SINTASSI DEL PROGRAMMA.

DATI

QUANDO SI ESEGUE UN PROGRAMMA LA MEMORIA
DALL'ELABORATORE DEVE CONTENERE:

- IL PROGRAMMA
- I DATI

I DATI VANNO PERO' LETTI DALLE UNITA' PERIFERICHE

PER ESEGUIRE CORRETTAMENTE QUESTA OPERAZIONE SI DEVE
CONOSCERE COME I DATI SONO FISICAMENTE MEMORIZZATI
SU TALI DISPOSITIVI

QUESTO OBBLIGA A CONOSCENZE INFORMATICHE CHE SONO DEL
TUTTO ESTRANEE ALLA CULTURA DELL'UTILIZZATORE FINALE

UN LINGUAGGIO OTTIMALE DOVREBBE ESSERE QUINDI
INDIPENDENTE DAL MODO IN CUI I DATI SONO REGISTRATI
SULLE PERIFERICHE

L'UTILIZZATORE FINALE POTREBBE IN TAL MODO CONCENTRARSI
ESCLUSIVAMENTE SUL PROBLEMA DA RISOLVERE

LINGUAGGI

ANZICHE' COMPIERE UNA TRADUZIONE VERA E PROPRIA,
PER ALCUNI LINGUAGGI HA LUOGO UNA "INTERPRETAZIONE"
CHE AVVIENE COMANDO PER COMANDO.

IN TAL CASO NON SI PARLA PIU' DI "COMPILATORE" MA DI
"INTERPRETE".

I PRINCIPALI LINGUAGGI OGGI DISPONIBILI SONO:

- FORTRAN (PER APPLICAZIONI SCIENTIFICHE)
- COBOL (PER APPLICAZIONI COMMERCIALI)
- PL/1 (CON LE CARATTERISTICHE SIA DEL FORTRAN
CHE DEL COBOL)
- BASIC (USATO SOPRATTUTTO SUI MICRO E PERSONAL)
- SAS (EVOLUZIONE DEL PL/1)

ESEMPIO PROGRAMMA

PROBLEMA : FARE LA SOMMA DI DUE NUMERI

SOLUZIONE CON UN IPOTETICO LINGUAGGIO NATURALE :

LEGGI PRIMO NUMERO
LEGGI SECONDO NUMERO
SOMMA I DUE NUMERI
STAMPA IL RISULTATO
FINE

PROGRAMMA IN LINGUAGGIO AD ALTO LIVELLO (FORTRAN) :

```
READ NUM1  
FORMAT (.....)  
READ NUM2  
FORMAT (.....)  
  
TOTALE = NUM1 + NUM2  
  
PRINT TOTALE  
FORMAT (.....)  
STOP
```

ESEMPIO PROGRAMMA

RIPRENDIAMO AD ESEMPIO IL PROGRAMMA VISTO IN
PRECEDENZA. ESSO POTREBBE ESSERE COSI' RIDOTTO:

LEGGI DATI

TOTALE = NUM1 + NUM2

PRINT TOTALE

STOP

CON I SEGUENTI VANTAGGI:

- DIMINUZIONE DEL NUMERO DI ISTRUZIONI
- MAGGIORE CHIAREZZA DEL PROGRAMMA
- ASSOLUTA INDIPENDENZA DAL MODO DI REGISTRAZIONE
DEI DATI

12

PRESENTAZIONE

DEL SAS

QUANDO OTTENIAMO I RISULTATI?

IL SAS PUO' LAVORARE IN DUE MODI:

- BATCH

- INTERATTIVO

STATISTICAL ANALYSIS SYSTEM

Cos'É IL S.A.S. ?

IL S.A.S. É UN SISTEMA SOFTWARE CHE FORNISCE I NECESSARI STRUMENTI PER ANALIZZARE DATI.

SI DEFINISCE SISTEMA SOFTWARE IN QUANTO É COMPOSTO DA:

UN LINGUAGGIO PER LA MANIPOLAZIONE DEI DATI

+

UNA LIBRERIA DI PROCEDURE PRE-CONFEZIONATE PER USO GENERALE

COME SI USA IL SAS ?

SCRIVENDO PROGRAMMI

IL LINGUAGGIO SAS, ESTREMAMENTE POTENTE, AVANZATO
È COMPLETO PERMETTE LA SCRITTURA DI PROGRAMMI SIN-
TETICI, MA TAGLIATI SU MISURA

SFRUTTANDO LE PROCEDURE DI LIBRERIA

SI POSSONO RICHIAMARE LE VARIE PROCEDURE SIA
DALL'INTERNO DEI PROGRAMMI SIA INTERATTIVAMENTE
DA TERMINALE

MODO BATCH

IL NOSTRO PROGRAMMA VERRA' SOTTOPOSTO ALL'ELABORATORE
INSIEME AI PROGRAMMI DI ALTRI UTILIZZATORI.

QUANDO L'ELABORATORE AVRA' ESAMINATO TUTTI I LAVORI
FORNIRA' I DIVERSI RISULTATI CHE VERRANNO DIVISI E
DISTRIBUITI AI VARI UTENTI.

QUESTO MODO DI LAVORARE E' TIPICO DEI SISTEMI OPERATIVI
DOS - OS - MVS.

E' POSSIBILE OPERARE IN QUESTO MODO ANCHE CON I SISTEMI
INTERATTIVI TSO - ICCF - VM/CMS.

MODO INTERATTIVO

IL NOSTRO PROBLEMA VIENE DIGITATO SU DI UN TERMINALE
PASSO PER PASSO. IL SAS CONTROLLA LA CORRETTEZZA DI
QUELLO CHE VIENE SCRITTO E AD OGNI FINE PASSO FORNISCE
I RISULTATI PARZIALI RELATIVI AI PASSI INTRODOTTI.

SI PUO' OPTARE PER QUESTO MODO DI LAVORARE SOLO SE
ABBIAMO UN TERMINALE E UN SISTEMA INTERATTIVO
COME IL TSO, ICCF, VM/CMS.

UN PASSO VIENE COSIDERATO CHIUSO QUANDO:

- INIZIA UN NUOVO PASSO (CON LE PAROLE
DATA O PROC(EDURE)).
- SI INCONTRA LA PAROLA "RUN;"
- SI INCONTRA LA FINE SESSIONE (INDICATA CON
LA PAROLA "ENDSAS"
O CON /*{return})

Quali sono i vantaggi del SAS?

- FACILITA' DI UTILIZZO
- POTENZA
- RICCHEZZA DELLE LIBRERIE
- INTEGRAZIONE DI VARI STRUMENTI E DISCIPLINE

Inoltre

- LINGUAGGIO STRUTTURATO
- POSSIBILITA' DI UNIFICARE IL LINGUAGGIO DI PROGRAMMAZIONE DEI PROGRAMMATORI DI UN CENTRO DI CALCOLO (IL SAS SI PRESTA BENE A RISOLVERE I PROBLEMI PIU' DIVERSI)
- LINGUAGGIO PORTATILE E DIFFUSO

FACILITA' DI UTILIZZO

IL SAS E' STATO PROGETTATO PER ESSERE UTILIZZATO
ANCHE DA NON ESPERTI DI INFORMATICA CHE POSSONO
IMPARARE SOLO GLI ELEMENTI BASE CON POCO SFORZO
E POTER SFRUTTARE IL VANTAGGIO DI UTILIZZARE UN
ELABORATORE DI DATI.

POTENZA DEL SAS

IL SAS E' STATO PROGETTATO PER ESSERE UTILIZZATO
ANCHE DA ESPERTI DI INFORMATICA CHE POSSONO
SFRUTTARE QUASI TUTTE LE POSSIBILITA' OFFERTE DAL
SISTEMA OPERATIVO E DALL'ELABORATORE.

RICCHEZZA DELLE LIBRERIE

IL SAS E' DOTATO DI MOLTISSIME LIBRERIE CHE SPAZIANO
DALLA GRAFICA ALLA STATISTICA, ALLA RICERCA
OPERATIVA, E A MOLTE ALTRE DISCIPLINE.

ESISTONO INOLTRE TUTTA UNA SERIE DI PROCEDURE
SCRITTE DA UTENTI CHE SONO STATE MESSE A DISPOSIZIONE
DI TUTTI (VEDERE SUPPLEMENTAL LIBRARY E SUGI). QUASI
SEMPRE SI PUO' TROVARE LA SOLUZIONE DEL NOSTRO PROBLEMA
GIA' SVOLTA E SPESSO NON SOLO UNA SOLUZIONE.

INTEGRAZIONE DI VARIE DISCIPLINE

PER RISOLVERE CERTI PROBLEMI SPESSO SIAMO COSTRETTI AD AVVALERCI DI VARIE DISCIPLINE (STATISTICA GRAFICA ECONOMIA MATEMATICA ECC.): DI SOLITO I LINGUAGGI CHE SONO EFFICIENTI PER LA MATEMATICA NON LO SONO ALTRETTANTO PER LA GRAFICA O PER LA SCRITTURA DEI TESTI ECC. IL SAS INVECE E' STATO STUDIATO APPOSTA PER INTEGRARE NEL MIGLIORE DEI MODI PIU' PROBLEMATICHE ED POTER RISULTARE UNO STRUMENTO IL PIU' POSSIBILE UNIVERSALE.

CAPITOLO 1
CONSIDERAZIONI
GENERALI
SUL SAS

BREVI CENNI STORICI SUL SAS (Statistical Analysis System)

Nel 1966 è iniziata la realizzazione del SAS con il seguente scopo:

- 1) venire incontro a chi ha molti dati da analizzare e agevolargli i problemi connessi alla loro elaborazione.
- 2) permettere all'utente di concentrarsi SOLO sui risultati dell'elaborazione.

L'apprendimento del SAS si pone come alternativa alla necessità di conoscere:

- 1) Un linguaggio di programmazione (Cobol, Fortran, Pli, Basic o altro)
- 2) Programmi di utilità (Utility OS o UM ecc.)
- 3) Vari pacchetti di programmi (PACKAGES) specializzati nelle discipline necessarie allo svolgimento del nostro lavoro (Packges statistici, matematici, grafici, econometrici ecc.)

Il SAS nato per esigenze statistiche, si è poi sviluppato come SISTEMA DI USO GENERALE.

STATO DELL' ARTE DEL SAS HARDWARE

I sistemi sui quali gira attualmente il SAS sono.

IBM 370/303X/43XX/308X e compatibili

con i sistemi operativi:

OS/MVS - TSO

VM/CMS

DOS/VSE - ICCF

DIGITAL VAX 11/7XX con Sist. Operativo VMS

DATA GENERAL ECLIPSE serie MV con AOS/VS

PRIME serie 50 sotto PRIMOS

Futuri annunci sono:

SAS (ridotto?) sul Personal Computer IBM e compatibili
entro giugno '85.

(si dice che qualcuno lo ha già visto...)

70% PL/I

25% ASSEMBLER

Possibile 2.4% di scansio

STATO DELL' ARTE DEL SAS SOFTWARE

Gli attuali campi di applicazione del SAS sono:

SAS BASE - Analisi statistica, gestione archivi, produzione di tabulati, ecc.

GCHART }
GMAP } - SAS GRAPH - Produzione di carte tematiche, plots
GSLIDE } CARTE TEMATICHE tridimensionali a colori, ecc.

SAS ETS - Econometria, analisi delle serie temporali, ecc.

- APPLICAZIONE
CORSI }
- MAILING LIST X }
INTERESSE }
- STATISTICHE }
FREQUENZA }
SAS FSP - Procedure interattive per l'editing, data entry, Word processing, scrittura di lettere commerciali, gestione tabelle elettroniche, ecc.

SAS IMS-DL1 - Interfaccia per accedere ai data base IMS/VS o CICS/OS/VS.

SAS OR - Ricerca operativa e programmazione lineare.

MERRILL'S GUIDE - Valutazione delle prestazioni dei sistemi.

Futuri annunci sono:

DATA BASE?

- INTEGRAZIONE DATA BASES ESISTENTI,
- SYSTEM 2000
- APT

DOCUMENTAZIONE SAS

1) INTRODUCTORY GUIDE	principianti
2) USER'S GUIDE	manuale base (BLUONE)
3) PROGRAMMER'S GUIDE	esperti
4) APPLICATION GUIDE	problemi risolti
5) CMS COMPANION	problemi CMS
6) SAS GRAPH	SAS/GRAPH
7) SAS FSP	SAS/FSP
8) SAS ETS	SAS/ETS
9) SUPPLEMENTAL LIBRARY USER'S GUIDE	procedure utenti
10) SAS/IMS-DL/I	SAS/IMS
11) SAS/OR	SAS/OR
12) SUGI PROCEEDINGS	1981/1984
13) MERRILL'S GUIDE	performance
14) TECHNICAL REPORTS	per esperti
15) SAS COMMUNICATION	quadrimestrale

SUGI

Il SAS User Group International e' un'associazione senza fini di lucro a cui possono partecipare tutti gli utenti SAS senza alcun costo di iscrizione.

Dal 1976 viene organizzata una conferenza annuale negli Stati Uniti strutturata per permettere il massimo scambio di esperienze fra utenti, in tutti i settori di applicazione del SAS.

Esistono inoltre dei gruppi a livello nazionale del SUGI ed uno sara' attivato entro il primo quadrimestre 1985 anche in Italia.

CAPITOLO 2

FUNZIONAMENTO DEL SAS

COME SI PRESENTANO I PROBLEMI AL SAS ?

I PROBLEMI DEVONO ESSERE SUDDIVISI IN PASSI (STEPS)

IN SAS ESISTONO 2 CATEGORIE DI PASSI:

- DATA STEP
- PROCEDURE STEP

DATA

SI USANO PER CREARE FILES SAS PARTENDO DA ARCHIVI ESISTENTI (DI TIPO SAS O NON); I DATI LETTI POSSONO SUBIRE TRATTAMENTI PRIMA DI ESSERE RESTITUITI IN USCITA

PROC

SI USANO PER ANALIZZARE FILES SAS E PER LA PRODUZIONE DI TABULATI (RAPPORTI), STATISTICHE O ALTRI TIPI DI RISULTATO

(I RISULTATI POSSONO COSTITUIRE UN NUOVO ARCHIVIO DA ANALIZZARE DA UN'ALTRA PROC SUCCESSIVA)

funzionamento del SAS

I programmi SAS sono suddivisi in step.
Esistono due tipi di step:

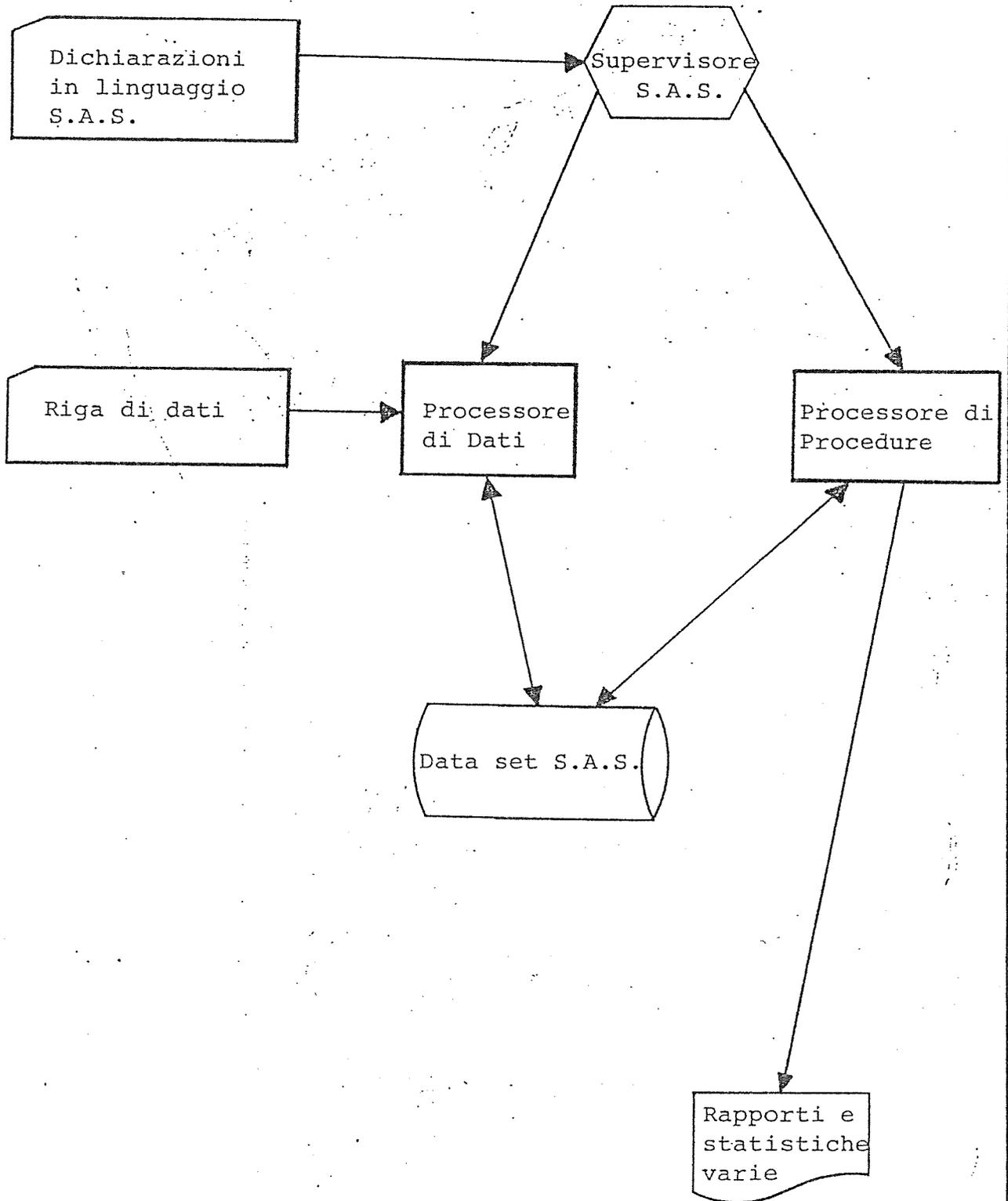
Step SAS

DATA

PROCEDURE

In generale gli step di data costruiscono files SAS partendo da files SAS o non e gli step PROC elaborano i files SAS.

SCHEMA GENERALE DI FUNZIONAMENTO DEL S.A.S.

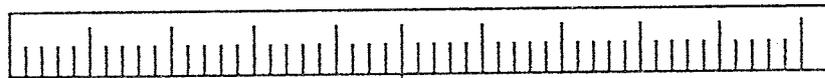


funzionamento del SAS

step DATA.

```
DATA SASDATA;  
  INPUT A B C D;  
CARDS;  
1 2 3 4  
5 6 7 8  
9 10 11 12  
;
```

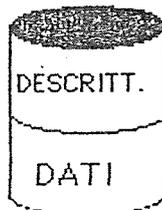
INPUT BUFFER



P D Y



FILE
SAS



funzionamento del SAS

COMPILAZIONE dello step DATA.

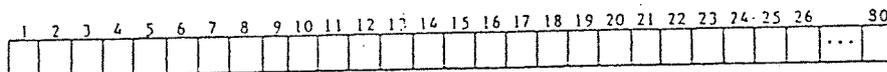
- Viene creato il **PDV** riservando lo spazio opportuno per le variabili, esattamente nell'ordine di apparizione nello step. I riferimenti a files SAS (istruz. SET, MERGE, ecc.) vengono considerati come riferimenti alle variabili costituenti il file SAS.
- Vengono creati i descrittori dei costruendi files SAS (tenendo conto di eventuali istruz. DROP o KEEP).

COMPILAZIONE DEL PASSO DI DATA

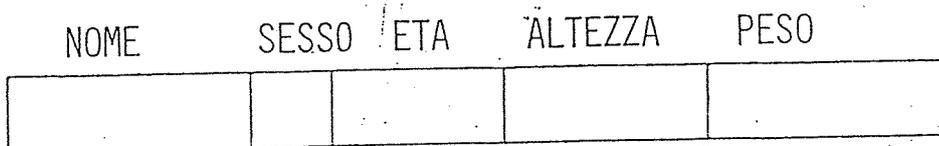
DATA SCUOLA;
INPUT NOME \$ 1-8 SESSO \$ 11 ETA. 13-14 ALTEZZA 16-18
PESO 21-22;
CARDS;

.....
.....
.....

INPUT BUFFER = AREA DI MEMORIA TEMPORANEA PER UN RECORD DI INPUT.



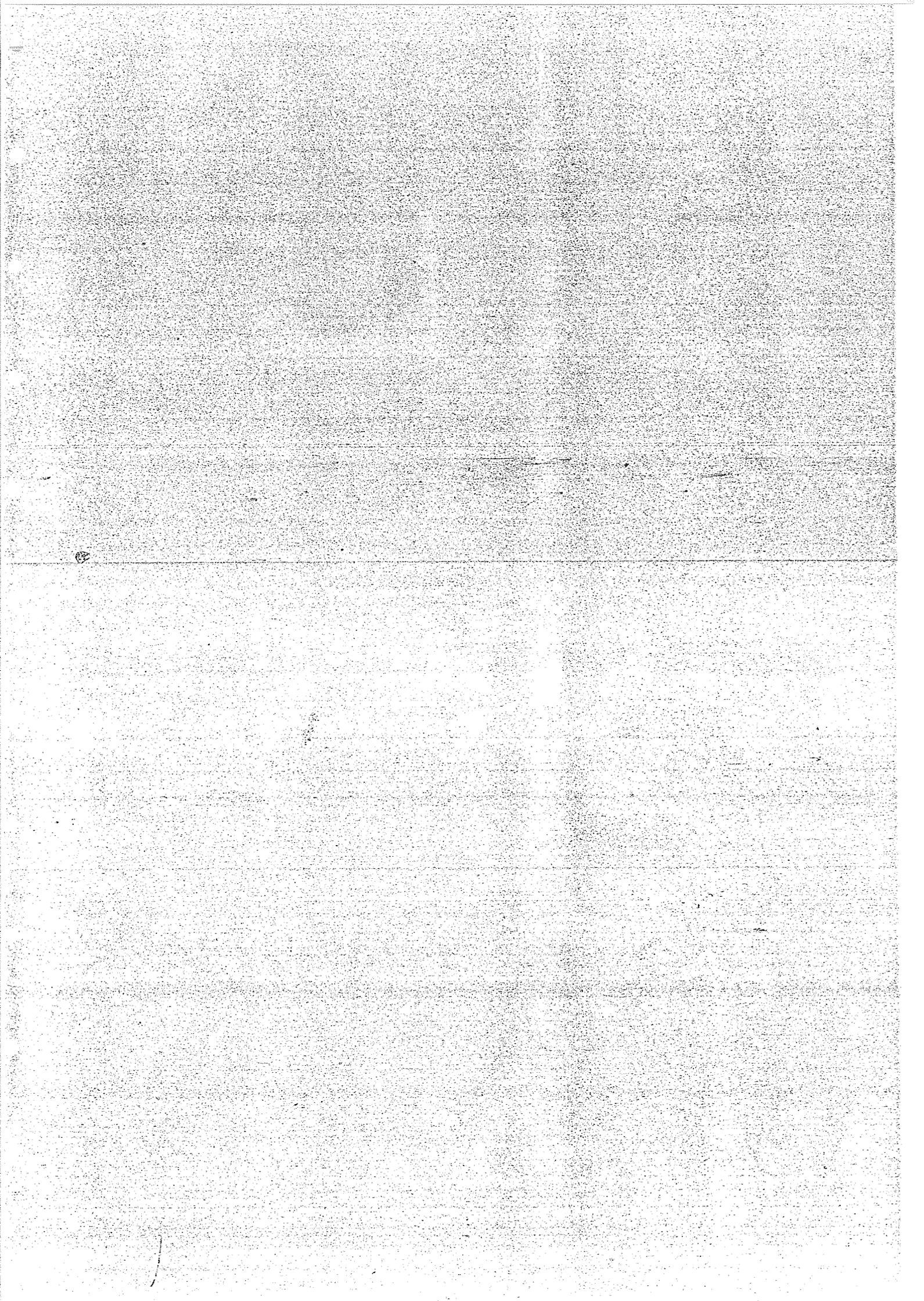
PROGRAM DATA VECTOR = AREA DI MEMORIA PER UNA OSSERVAZIONE SAS



LE INFORMAZIONI DESCRITTIVE VENGONO SCRITTE NEL DATA SET SAS
SAS data set WORK.SCUOLA.

work
disk
space

NOME SESSO ETA ALTEZZA PESO



funzionamento del SAS

ESECUZIONE dello step DATA.

- Le istruzioni di uno step vengono eseguite nell'ordine in cui si trovano a meno di salti, una dopo l'altra.
- Dopo l'esecuzione dell'ultima istruzione il SAS scrive i valori delle variabili del PDV nel dataset di uscita (solo quelle presenti nel descrittore del dataset) e assegna valore vuoto (missing value) a tutte le variabili del PDV (house keeping) e ritorna ad eseguire lo step iniziando con la prima istruzione eseguibile
- L'esecuzione si arresta quando i dati di ingresso sono terminati

ESECUZIONE DEL PASSO DI DATA

- IL PROGRAM DATA VECTOR È INIZIALIZZATO A MISSING (VALORE MANCANTE) PRIMA DI OGNI ESECUZIONE DEL PASSO DI DATA.
- IL PASSO DI DATA VIENE ESEGUITO UNA VOLTA PER OGNI OSSERVAZIONE IN INPUT.

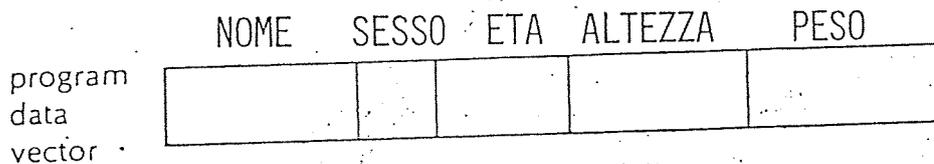
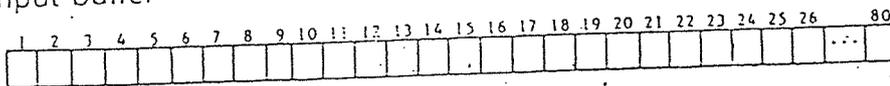
DATA SCUOLA;

INPUT NOME S 1-8 SESSO S 11 ETA 13-14
ALTEZZA 16-18 PESO 21-22;

CARDS;

GIULIO M	12	150	40
MARCO M	12	156	42
ELENA F	12	136	37

input buffer



SAS data set WORK.

work disk space

NOME SESSO ETA ALTEZZA PESO

ESEMPIO

CONSIDERIAMO IL SEGUENTE INSIEME DI DATI:

12345678901234567890123456789012345678901234567890 N° DI COLONNA

ANTONUCCI	M 12	140,5	035,6	1°	SOGGETTO
AZZOLLINA	M 13	143,0	038,0	2°	"
BARTALI	F 12	141,8	036,6	3°	"
::				
::				
ZUCHELLI	M 13	145,9	039,3	N°	"

CIASCUNA RIGA (OSSERVAZIONE) PUO' ESSERE PERFORATA SU DI UNA SCHEDA OPPURE ESSERE EDITATA IN UN ARCHIVIO (FILE O DATA SET) DA UNA PERFORATRICE O DA UN EDITORE

ESEMPIO DI UN PROGRAMMA S.A.S.

DATA CLASSE;

INPUT COGNOME &20, &11 SESSO &1, &13 ETA 2, &16 ALTEZZA 5,1
&22 PESO 5,1 ;

CARDS;

ANTONUCCI	M 12	140,5	035,6
AZZOLLINA	M 13	143,0	038,0
...		
ZUCHELLI	M 13	145,9	039,3

PROC PRINT;

PROC PLOT; PLOT PESO * ALTEZZA = SESSO;

REGOLE PER SCRIVERE UNA CORRETTA DICHIARAZIONE S.A.S.

LE DICHIARAZIONI S.A.S. :

- HANNO FORMATO LIBERO
- POSSONO CONTINUARE SU PIU' RIGHE
- POSSONO TROVARSI SU UNA RIGA IN NUMERO SUPERIORE AD UNO
- POSSONO AVERE I CAMPI SEPARATI DA PIU' CHE UN BLANK (SPAZIO)

E' CONSIGLIABILE SCRIVERE I PROGRAMMI IN MODO CHE DATA E PROC COMINCINO SEMPRE A COLONNA 1; IN QUESTO MODO SI EVIDENZIANO I PASSI DELLA PROCEDURA.

```
DATA ELENCO;  
  INPUT NOME $ ETA FUNZIONE $;  
  ANN_NASC=1984-ETA;  
  CARDS;  
GIOVANNI 54 A4  
ELIO 37 A7  
RENATO 42 A5  
;  
PROC PRINT;  
RUN;
```

COME NON CONVIENE SCRIVERE UN PROGRAMMA

```
DATA ELENCO;INPUT NOME $ ETA FUNZIONE $;ANN_NASC=  
1984-ETA;CARDS;  
GIOVANNI 54 A4  
ELIO 37 A7  
RENATO 42 A5  
PROC PRINT;  
/*
```

SUPERVISORE
S.A.S.

DATA STEP

PROCEDURE STEP

PREPARAZIONE DEL
DATA SET IN FOR-
MATO S.A.S.

ELABORAZIONE DEL
DATA SET S.A.S.

RIGHE DI INPUT

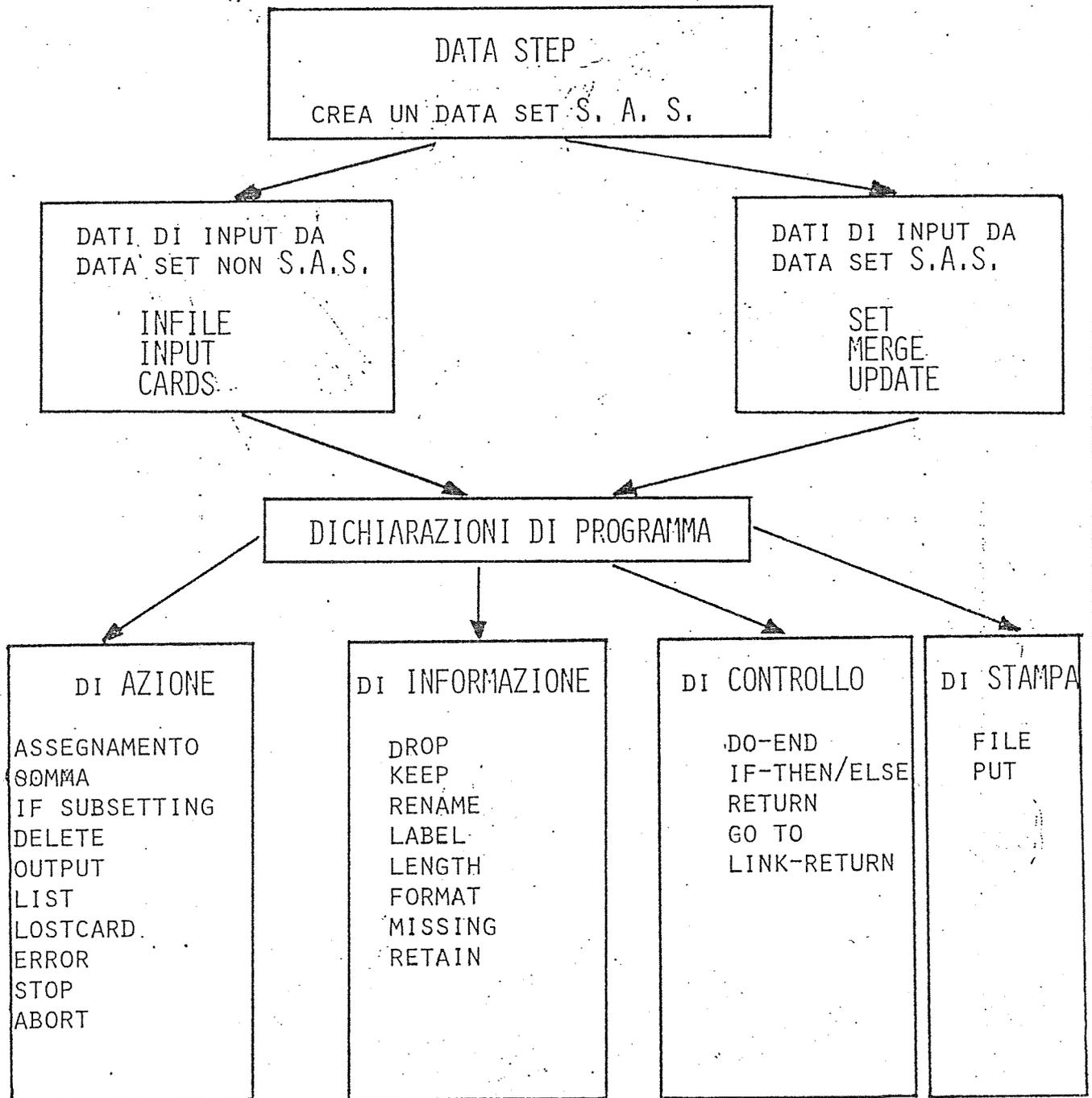
DATA;
INPUT X Y;
CARDS;
X1 Y1
X2 Y2

DATA SET

PROC MEANS;
VARS X Y;

GENERAZIONE DI
RAPPORTI E DI
STATISTICHE

COSA SERVE PER SCRIVERE UN PROGRAMMA S.A.S.



LE PROCEDURE S. A. S.

PER ELABORARE I DATI CONTENUTI IN UN DATA SET S.A.S. SI USANO

LE PROCEDURE, LE QUALI PERMETTONO DI:

- LEGGERE DATA SET
- ELABORARE I DATI (ES. FARE STATISTICHE)
- STAMPARE I RISULTATI
- CREARE NUOVI DATA SET

LE PROCEDURE PIU' NOTE POSSONO ESSERE RAGGRUPPATE NEL SEGUENTE MODO:

STATISTICA DESCRITTIVA

CORR
FREQ
MEANS
RANK
SUMMARY
UNIVARIATE

ANALISI E DISEGNO DI ESPERIMENTI

ANOVA
FUNCAT
GLM
NESTED
PROBIT
TTEST
VARCOMP

REGRESSIONE

GLM
NLIN
RSQUARE
STEPWISE
SYSREG

ANALISI MULTIVARIATA

CANCORR
CLUSTER
DISCRIM
FACTOR
GLM
GUTTMAN
NEIGHBOR
SCORE

ECONOMETRIA E ANALISI

DI SERIE TEMPORALI

AUTOREG
SPECTRA
SYSREG

GRAFICA

CHART
PLOT

SPECIALI

MATRIX
PLAN

INTERFACCIA

BMDP
CONVERT

UTILITIES S.A.S.

CONTENTS
COPY
DATASETS
DELETE
EDITOR
OPTIONS
PRINT
SORT
STANDARD

UTILITIES OS

PDS
PDSCOPY
PRINTTO
RELEASE
SOURCE
TAPELABEL

CAPITOLO 3

DATA STEP

DOPO LA FASE DI CREAZIONE DEL DATA SET, L'ELABORATORE SCRIVE:

NOTE: DATA SET WORK.CLASSE HAS NN OBSERVATIONS AND MM
VARIABLES xxxx OBS/TRK

IN QUESTO MODO IL SISTEMA SEGNALE CHE IL DATA SET WORK.CLASSE
È UN DATA SET TEMPORANEO E, AL TERMINE DEL LAVORO IL SISTEMA
LO CANCELLERÀ DAL DISCO.

PRINCIPALI DICHIARAZIONI DEL PASSO DATA

INPUT - CARDS

ASSEGNAZIONE

TOTALIZZAZIONE

RETAIN

SUM

IF-THEN-ELSE

DELETE

IF SUBSETTING

OUTPUT

ASSEGNAZIONE - DROP - KEEP

DO

ALCUNE FUNZIONI (MEAN, ABS, SUM,
SUBSTR, INDEX)

DICHIARAZIONE INPUT

ESISTONO 3 MODI DI ESEGUIRE UNA DICHIARAZIONE INPUT:

- A LISTA
- PER COLONNA
- FORMATTATA

LISTA

- LE VARIABILI DA LEGGERE COMPAIONO NELLA SCHEDA INPUT NELLO STESSO ORDINE CON CUI ESSE APPAIONO NEI DATI
- IL S.A.S. RICERCA NELLA RIGA IL PRIMO CARATTERE NON BIANCO E DA QUEL PUNTO IN POI, FINO AL PROSSIMO BIANCO (BLANK), IL CAMPO LETTO RAPPRESENTA IL VALORE DELLA VARIABILE

VANTAGGI DI UNA DICHIARAZIONE INPUT A LISTA

- FACILITÀ DI SCRITTURA
- NON SIAMO COSTRETTI A CONTARE LE COLONNE

REGOLE DI SCRITTURA

- \$ È USATO PER INDICARE UNA VARIABILE CARATTERE
- OGNI CAMPO DEVE ESSERE SPECIFICATO IN ORDINE
- I CAMPI DEVONO ESSERE SEPARATI DA UNO O PIU' SPAZI
- LE VARIABILI CARATTERE POSSONO ESSERE AL MASSIMO DI 8 CARATTERI E NON DEVONO AVERE SPAZI COME CARATTERI PROPRI DEL CAMPO (ESEMPIO: ROSSI E, VIENE INTERPRETATO COME ROSSI E, E NON COME UN UNICO CAMPO)

DICHIARAZIONE CARDS

SPECIFICA CHE I DATI PER IL PROGRAMMA SEGUONO QUESTA
SCHEDA.

LA TERMINAZIONE DEI DATI È INDICATA DA UN ";" A COLONNA 1.

SE NEI DATI È PRESENTE UN ";" SI USA LA DICHIARAZIONE

CARDS4 LA QUALE ASSUME COME FINE DAI DATI "::::" DA COLON-

NA 1 A COLONNA 4

ESERCIZIO

SCRIVERE UN PROGRAMMA CHE LEGGA DEI DATI MENSILI RELATIVI ALLA
CADUTA DI PIOGGIA IN UNA DATA ZONA E:

- 1) CREI UN APPOSITO DATA SET
- 2) STAMPI IL DATA SET
- 3) DETERMINI LA MEDIA DI CADUTA ANNUALE

1) CREAZIONE:

DATA PIOGGIA;

INPUT MESE 83, PIOGGIA 6.1;

CARDS;

GEN 0200,5
FEB 0236,7
MAR 0184,2
APR 0083,2
MAG 0084,2
GIU 0054,2
LUG 0005,6
AGO 0009,3
SET 0021,8
OTT 0123,5
NOV 0167,6
DIC 0324,7

2) PROC PRINT DATA=PIOGGIA;

DICHIARAZIONE INFILE

L'INFILE IDENTIFICA GLI ARCHIVI "NON SAS" DAI QUALI
VOGLIAMO CHE SIANO LETTE LE INFORMAZIONI.
GLI ARCHIVI "NON SAS" SONO TUTTI QUEGLI ARCHIVI GESTITI
E CONOSCIUTI DAL SISTEMA OPERATIVO E NON NECESSARIAMENTE
CREATI DA UN PROGRAMMA SAS.

INFILE FILEMIO;

FILEMIO È UN NOME CHE IDENTIFICA UNA SCHEDA CONTROLLO
DEL SISTEMA OPERATIVO LA QUALE A SUA VOLTA INDIVIDUA
IL FILE DESIDERATO.

INFILE FILE001;

RICHIAMA LE SCHEDE DOS (SE QUESTO È IL SISTEMA OPERATIVO
CHE STIAMO USANDO)

//DLBL FILE001,'NOME.DEL.DATASET'

//EXTENT

//ASSGN

O LA SCHEDA OS (SE USIAMO L'OS)

//FILE001 DD DSNAME=...

O LA DICHIARAZIONE CMS

FILEDEF FILE001 DISK

DICHIARAZIONE INFILE (PARAMETRI PRINCIPALI)

POICHÈ CON LA DICHIARAZIONE INFILE SI DEFINISCONO
FILES "NON SAS" E QUINDI SCONOSCIUTI AL SAS SI
DEBONO FORNIRE CON QUESTA DICHIARAZIONE TUTTE LE
INFORMAZIONI NECESSARIE ALLA ELABORAZIONE DEL FILE
CHE NON POSSONO ESSERE DEDOTTE O RITROVATE SULLE
SCHEDE CONTROLLO (IN DOS NON È POSSIBILE DEFINIRE
IL FILE NELLE SCHEDE CONTROLLO)

INFILE PIPPO VSAM;

IL FILE PIPPO È DI TIPO VSAM E TUTTE LE INFORMAZIONI
RISIEDONO NELL'APPOSITO CATALOGO;

INFILE LIBRO1(MEMBRO1);

IL FILE È COSTITUITO DA UN MEMBRO DI LIBRERIA;

INFILE FILE003 RECFM=FB LRECL=80 BLKSIZE=800;

SI FORNISCONO INFORMAZIONI RIGUARDO AL FORMATO DELLA
REGISTRAZIONE, LA LUNGHEZZA DELLA REGISTRAZIONE LOGICA E
QUELLA FISICA

INFILE FILETUO EOF=LABEL2;

.....
LABEL2:

LABEL2 È L'ETICHETTA DELLA ISTRUZIONE ALLA QUALE
IL SAS DEVE SALTARE A FINE LETTURA DI FILETUO;

MISSING VALUES (VALORI VUOTI)

IN MOLTI CASI, I DATI DI INGRESSO INCLUDONO VALORI VUOTI O MISSING VALUES (ES. RISPOSTE MANCANTI AD UNA DOMANDA DI UN QUESTIONARIO);

VALORI VUOTI ASSEGNATI A VARIABILI NUMERICHE:

- LA PRESENZA DI UN VALORE VUOTO PER UNA VARIABILE NUMERICA SI RAPPRESENTA CON UN PUNTO NEL CAMPO DELLA VARIABILE (INPUT A LISTA) O CON UN BLANK (INPUT A COLONNA/FORMATTATO),
- INTERNAMENTE IL SAS ASSEGNA UN CODICE PARTICOLARE AL VALORE VUOTO, E IN STAMPA VIENE RAPPRESENTATO CON UN PUNTO (,)

VALORI VUOTI ASSEGNATI A VARIABILI ALFABETICHE:

- LA PRESENZA DI UN VALORE VUOTO PER UNA VARIABILE ALFANUMERICA SI RAPPRESENTA CON UN PUNTO NEL CAMPO DELLA VARIABILE (INPUT A LISTA) O CON UN BLANK (INPUT COLONNA/FORMATTATO)
- INTERNAMENTE IL SAS ASSEGNA UN CODICE PARTICOLARE AL VALORE VUOTO, E IN STAMPA VIENE RAPPRESENTATO CON BLANK

PER VEDERE SE IL VALORE DI UNA VARIABILE È VUOTO SI USANO I SEGUENTI IF:

VARIABILE NUMERICA: IF ALTEZZA=, THEN DELETE;

VARIABILE ALFABETICA: IF SESSO=' ' THEN DELETE;

LE PROC TRATTANO I VALORI VUOTI IN VARI MODI, E MOLTE DI ESSE ESCLUDONO TALI VALORI PRIMA DALL'ELABORAZIONE

CONCETTO DI DATA SET DI TIPO S.A.S.

UN DATA SET DI TIPO S.A.S. É UN INSIEME DI DATI OMOGENEI ORGANIZZATI IN FORMA RETTANGOLARE.

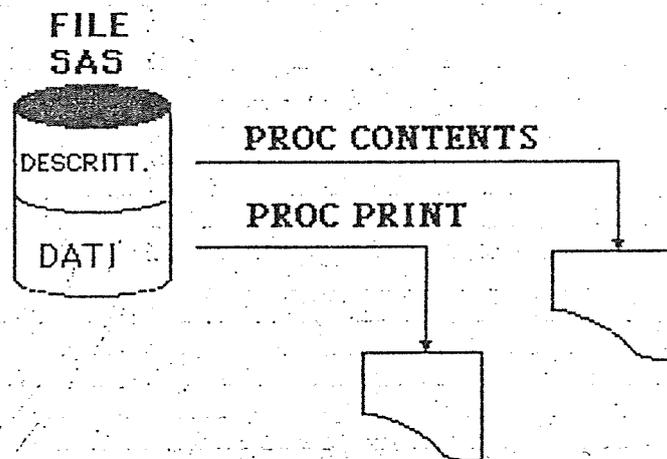
ESEMPIO:

	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0
Osservazione 1	ANTONUCCI										M	140,5	035,6																	
Osservazione 2	AZZOLLINA										M	143,0	038,0																	
Osservazione 3	BARTALI										F	141,8	036,6																	
.....																	
.....																	
Osservazione N	ZUCHELLI										M	145,9	039,3																	

- LE COLONNE SONO CHIAMATE VARIABILI
- CIASCUNA VARIABILE HA UN NOME
- CIASCUN NOME DEVE SEGUIRE LE SEGUENTI REGOLE SINTATTICHE:
 - DEVE ESSERE LUNGO DA 1 A MAX 8 CARATTERI
 - DEVE COMINCIARE CON CARATTERI DA A + Z OPPURE CON UN _ (SOTTOLINEATURA)
 - PUO' CONTENERE CIFRE AL SUO INTERNO
 - SI CONSIGLIA DI ASSEGNARE SEMPRE NOMI MNEMONICI

funzionamento del SAS

FILES SAS.



T I P I D I V A R I A B I L I

UNA VARIABILE PUO' ESSERE DI DUE TIPI:

- NUMERICA (Es. ETA, ALTEZZA, PESO)
- CARATTERE (Es. COGNOME, SESSO)

IL NUMERO MAX DI VARIABILI USABILI IN UN PROGRAMMA S.A.S. É 4000.

IN QUANTO AL NUMERO DI OSSERVAZIONI GESTIBILI DA UN PROGRAMMA S.A.S. ESO DIPENDE DALLA DISPONIBILITÀ DI SPAZIO DISCO NELLA CONFIGURAZIONE DELL'ELABORATORE USATO.

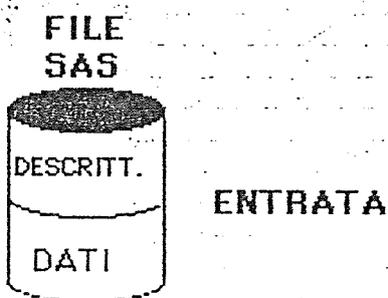
- UNA VARIABILE E' CARATTERIZZATA DALL' AVERE
UNA LUNGHEZZA:
 - LE VARIABILI NUMERICHE SONO LUNGHE 8 BYTES
 - LE VARIABILI CARATTERI POSSONO ASSUMERE
UNA LUNGHEZZA COMPRESA TRA 1 E 200 BYTES
(CARATTERI)

- LE VARIABILI ASSUMONO VALORI CHE DIPENDONO DAL-
L'ELABORAZIONE CHE SI STA' EFFETTUANDO. IN
PARTICOLARI SITUAZIONI NON ESISTONO DATI
ASSOCIABILI AD UNA VARIABILE (LA VARIABILE
CONTIENE IL VALORE VUOTO O "MISSING VALUE")

funzionamento del SAS

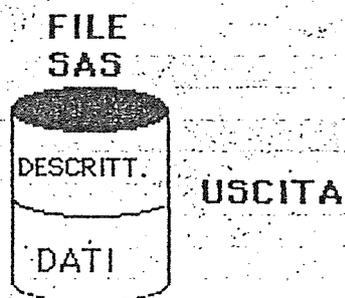
Da File SAS a File ~~non~~ SAS.

DATA USCITA;
SET ENTRATA;



P D V

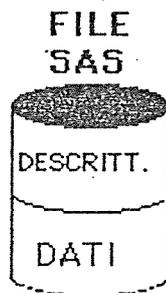
A	B	C	D	_ERROR_	_N_
---	---	---	---	---------	-----



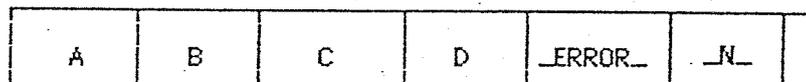
funzionamento del SAS

Da File SAS a File non SAS.

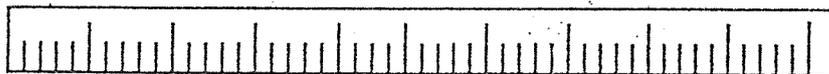
```
DATA _NULL_;  
  SET SASDATA;  
  FILE USCITA;  
  PUT A B C D;
```



P D V



OUTPUT BUFFER



ISTRUZIONI DI ASSEGNAMENTO

LE DICHIARAZIONI DI ASSEGNAMENTO SONO USATE PER CREARE NUOVE VARIABILI E PER MODIFICARNE ALTRE ESISTENTI.

ESEMPIO: VARIABILE=ESPRESSIONE

V2=VAR1;

ESPRESS=A+B/CxD;

DIVIS=NUMER/6 ;

DATA ARCHIVIO;

INPUT POP1982 POP1983;

VARIAZ=POP1983-POP1982;

CARDS;

50657343 52789543

ALLA VARIABILE VARIAZ IL S.A.S. ASSEGNERÀ IL VALORE DATO DALLA DIFFERENZA TRA LA VARIABILE POP1983 E POP1982.

L'ARCHIVIO ARCHIVIO CONTERRÀ LE VARIABILI POP1982 POP1983 VARIAZ

E' POSSIBILE ASSEGNARE COSTANTI NUMERICHE O ALFANUMERICHE:

ESEMPIO:

INDICE=5;

DESCRIZ='PIOGGIA' ;

IL S.A.S. METTE A DISPOSIZIONE UNA SERIE DI FUNZIONI MATEMATICHE MEDIANTE LE QUALI É POSSIBILE ASSEGNARE UN VALORE AD UNA VARIABILE:

TOTALIZZAZIONE CON SUM FUNCTION

PER EVITARE I PROBLEMI INDOTTI DALLA PRESENZA DI VALORI
"MISSING" SI FA USO DELLA FUNZIONE SUM:

ESEMPIO:

DATA SOMMA;

RETAIN TOTALE;0;

INPUT PARZIALE;

TOTALE=SUM(TOTALE,PARZIALE);

CARDS;

10

30

34

ECC, ECC,

PROC PRINT;

OBS	TOTALE	PARZIALE
1	10	10
2	40	30
3	40	
4	74	34
ECC,	ECC,	ECC,

TOTALIZZAZIONE CON SUM STATEMENT

E' POSSIBILE OMETTERE LA RETAIN USANDO LA SUM STATEMENT

ESEMPIO:

```
DATA SOMMA;  
INPUT PARZIALE;  
      TOTALE+PARZIALE;  
CARDS;  
  10  
  30  
  .  
  34  
  ECC.  
  
PROC PRINT;
```

OBS	PARZIALE	TOTALE
1	10	10
2	30	40
3	.	40
ECC.	ECC.	ECC.

IF-THEN-ELSE

In SAS e' possibile condizionare l'esecuzione di una istruzione all'avverarsi o meno di una condizione usando l'istruzione IF.

Sintassi:

```
IF <condizione>  
  THEN <dichiarazione eseguibile>;  
  ELSE <dichiarazione eseguibile>;
```

Oppure:

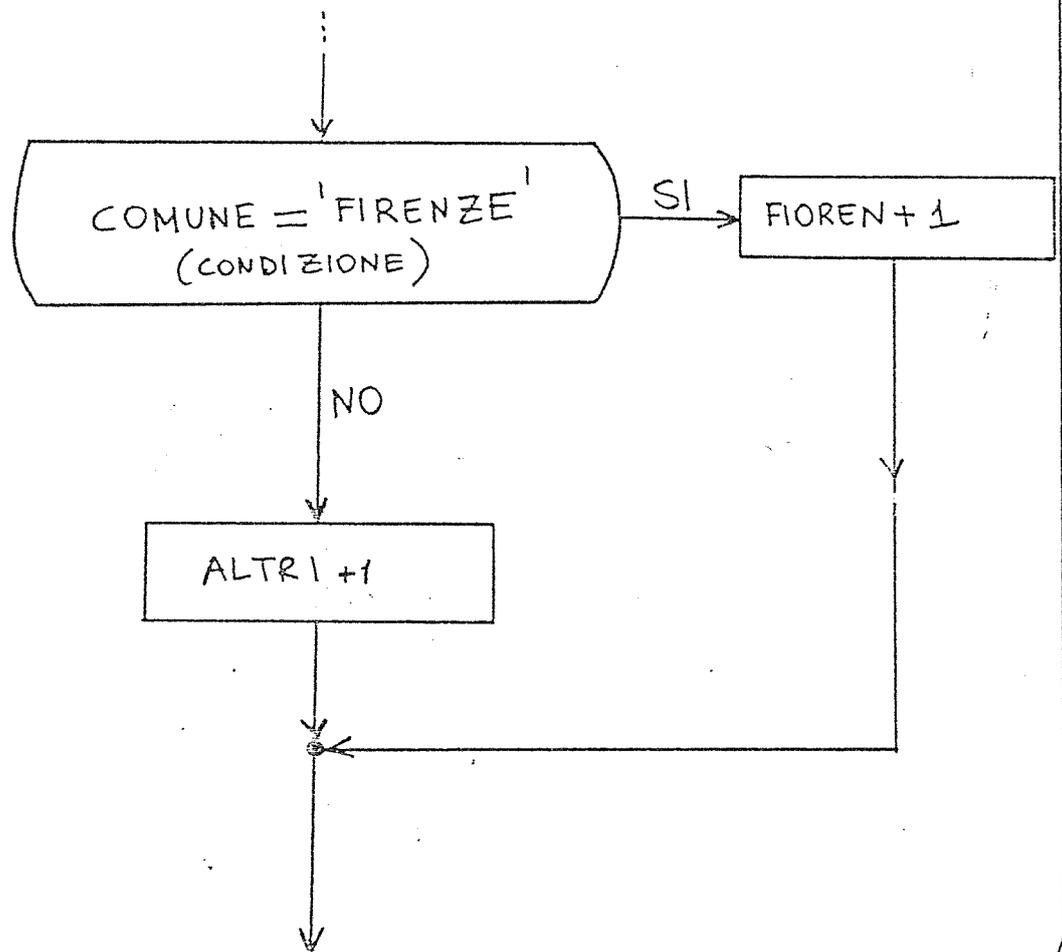
```
IF <condizione>  
  THEN <dichiarazione eseguibile>;
```

ESEMPIO:

```
IF COMUNE = 'FIRENZE' THEN FIOREN+1;  
ELSE ALTRI+1;
```

QUESTA DICHIARAZIONE IF DIFFERISCE DALLA PRIMA PER
LA PRESENZA DEGLI OPERATORI THEN/ELSE ,

DIAGRAMMA DI FUNZIONAMENTO DELLA IF



NELLE ESPRESSIONI LOGICHE (CONFRONTI) POSSONO ESSERE USATI

SEI OPERATORI :

FORMA 1	FORMA 2	SIGNIFICATO
<	LT	MINORE DI (LESS THAN)
>	GT	MAGGIORE DI (GREATER THAN)
=	EQ	UGUALE A (EQUAL THAN)
<=	LE	MINORE O UGUALE A (LESS THAN OR EQUAL TO)
>=	GE	MAGGIORE O UGUALE A (GREATER THAN OR EQUAL TO)
≠	NE	DIVERSO DA (NOT EQUAL TO)

GLI OPERATORI POSSONO ESSERE SCRITTI INDIFFERENTEMENTE NELLA FORMA 1 o 2 (NON TUTTI I TERMINALI O LE STAMPATRICI METTONO A DISPOSIZIONE I CARATTERI NECESSARI ALLA FORMA 1)

ESEMPIO:

IF ETA < 18 THEN RAGAZZI+1;

IF 18 <= ETA LT 75 THEN ADULTI+1;

IF ETA GE 75 THEN MATURI+1;

Istruzione DELETE

In SAS e' possibile controllare quali osservazione vogliamo che siano trascritte nel data set sas di uscita con l'istruzione DELETE.

Sintassi:

```
DELETE;
```

Esempio:

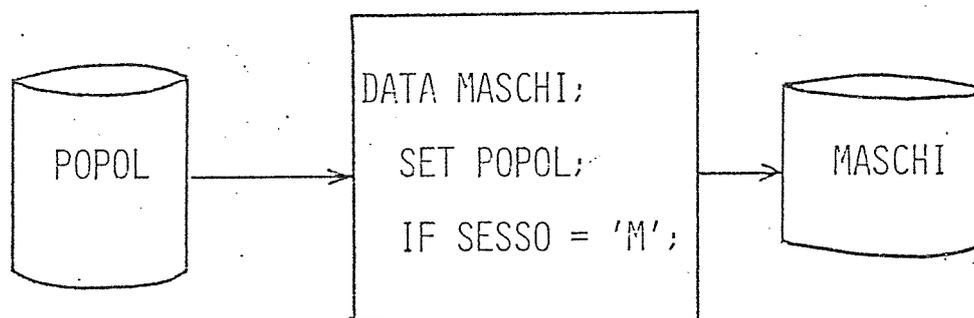
```
IF SESSO='M'  
  THEN DELETE;  
  ELSE FEMMINE+1;
```

L'esecuzione di una DELETE provoca il trasferimento della esecuzione all'inizio dello step dopo una re-inizializzazione del PDU (House-keeping).

DICHIARAZIONE IF-SUBSETTING

UN MODO SEMPLICE DI INTRODURRE AL FUNZIONAMENTO DI UNA DICHIARAZIONE IF ^{SUBSETTING} È IL SEGUENTE:

SUPPONIAMO DI AVERE A DISPOSIZIONE UN ARCHIVIO DELLA POPOLAZIONE E DI VOLER SELEZIONARE DA ESSO TUTTI GLI INDIVIDUI MASCHI.



LA DICHIARAZIONE IF SESSO='M'; È CHIAMATA 'SUBSETTING IF';

IL MODO DI LAVORARE DI QUESTA DICHIARAZIONE RISULTA SIMILE AD UNA PORTA LA QUALE PERMETTE AD UN'OSSERVAZIONE DI "PASSARE" SOLO SE LA CONDIZIONE DELL'IF (SESSO='M') RISULTA VERA.

NEL CASO CHE ESSA RISULTI FALSA (IL SOGGETTO È UNA FEMMINA) ALLORA L'OSSERVAZIONE NON VIENE INSERITA NEL NUOVO ARCHIVIO MASCHI.

ESEMPI

```
DATA MASCHI;  
  SET POPOL;  
  IF SESSO <'M' THEN DELETE;
```

Si ottiene lo stesso risultato scrivendo:

```
DATA MASCHI;  
  SET POPOL;  
  IF SESSO='M';
```

IF-subsetting con condizione composta:

```
DATA DONNE_SP;  
  SET POPOL;  
  IF SESSO='F' AND STATO_CU='C';
```

Dichiarazioni DROP e KEEP

Tutte le variabili incontrate nello step di DATA sono presenti nel PDU.

Per controllare le variabili che saranno inserite nel data set SAS si usano la DROP e la KEEP.

In assenza di DROP/KEEP tutte le variabili del PDU sono aggiunte al data set SAS.

Sintassi:

DROP lista di variabili;

KEEP lista di variabili;

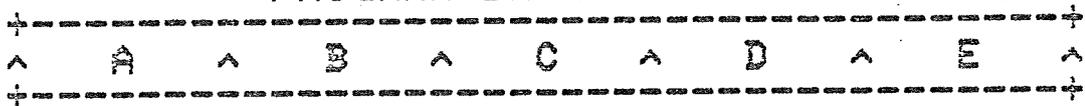
La DROP indica le variabili che NON DEVONO essere presenti nel data set SAS.

La KEEP indica le variabili che DEVONO essere presenti nel data set SAS.

ESEMPI

```
DATA ESEMPIO;  
  DROP A B C S;  
  D=A*B;  
  E=C;  
  CARDS;  
10 10 COSTANT1  
10 20 COSTANT2  
10 30 COSTANT3  
.....  
.....  
PROC PRINT;  
RUN;
```

PROGRAM DATA VECTOR



DATA SET A		
ENTRATE	USCITE	SALDO
2000000	1000000	1000000
3000000	500000	2500000

ESEMPIO:

DATA B;

SET A;

DROP ENTRATE USCITE;

VEETTORE DEI DATI

ENTRATE	USCITE	SALDO
---------	--------	-------

DATA SET B
SALDO
1000000
2500000

DA RIPORTARE NEL DATASET B

Dichiarazione DO iterativo

E' possibile far ripetere al SAS l'esecuzione di un gruppo di dichiarazioni tramite la DO.

Sintassi:

```
DO <indice> = <start> TO <stop> BY <step>;
```

```
DO <indice> = <start> TO <stop>  
  BY <step> WHILE <condiz.>,...;  
  UNTIL <condiz.>,...;
```

ESEMPIO:

```
DO I = 1 TO 100;  
  QUAD=I**2;  
  CUBO=I**3;  
  QUAD=I**4;  
END;
```

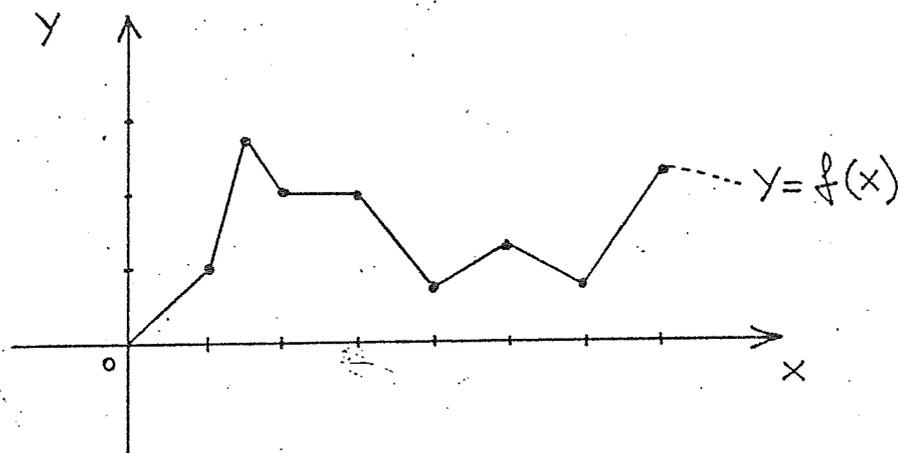
```
DO WHILE ( I>100 );  
  <Dichiaraz1>;  
  <Dichiaraz2>;  
  .....  
  <DichiarazN>;  
END;
```

```
DO I = 1,7,35;  
  <dichiaraz1>;  
  .....  
  <dichiarazN>;  
END;
```

```
DO UNTIL ( I>100 );  
  <Dichiaraz1>;  
  <Dichiaraz2>;  
  .....  
  <DichiarazN>;  
END;
```

PROCEDURA PLOT

LA NECESSITÀ DI INDIVIDUARE L'EVOLVERSI (L'ANDAMENTO, IL TREND) DI UN FENOMENO (ES. NASCITE, CRISI, ELEZIONI, ECC.) CI PORTA IN MOLTE OCCASIONI A DISEGNARE DEI GRAFICI DEL TIPO:



AL FINE DI OTTENERE DISEGNI DEL GENERE, IL SAS METTE A DISPOSIZIONE LA PROCEDURA PLOT.

SINTASSI:

```
PROC PLOT OPZIONI DATA=NOME ARCHIVIO SAS;
```

ESEMPIO D'USO DELLA PROC PLOT

DESIDERIAMO TRACCIARE IL GRAFICO DELLA SEGUENTE EQUAZIONE (RETTA):

$$Y = 2 X + 25$$

FACENDO VARIARE LA VARIABILE INDIPENDENTE X DA 0 A 100 SECONDO INCREMENTI COSTANTI DI 5.

IL PROGRAMMA È IL SEGUENTE:

```

A ← { DATA PUNTI;
      DO X=0 TO 100 BY 5;
      Y=2*X+25;
      OUTPUT;
      END;

B ← { PROC PLOT;
      PLOT Y*X='x';
      TITLE GRAFICO DI Y=2X+25;
    }

```

GRAFICO DI $Y=2X+25$

PLOT OF $Y=2X$ SYMBOL USED IS *

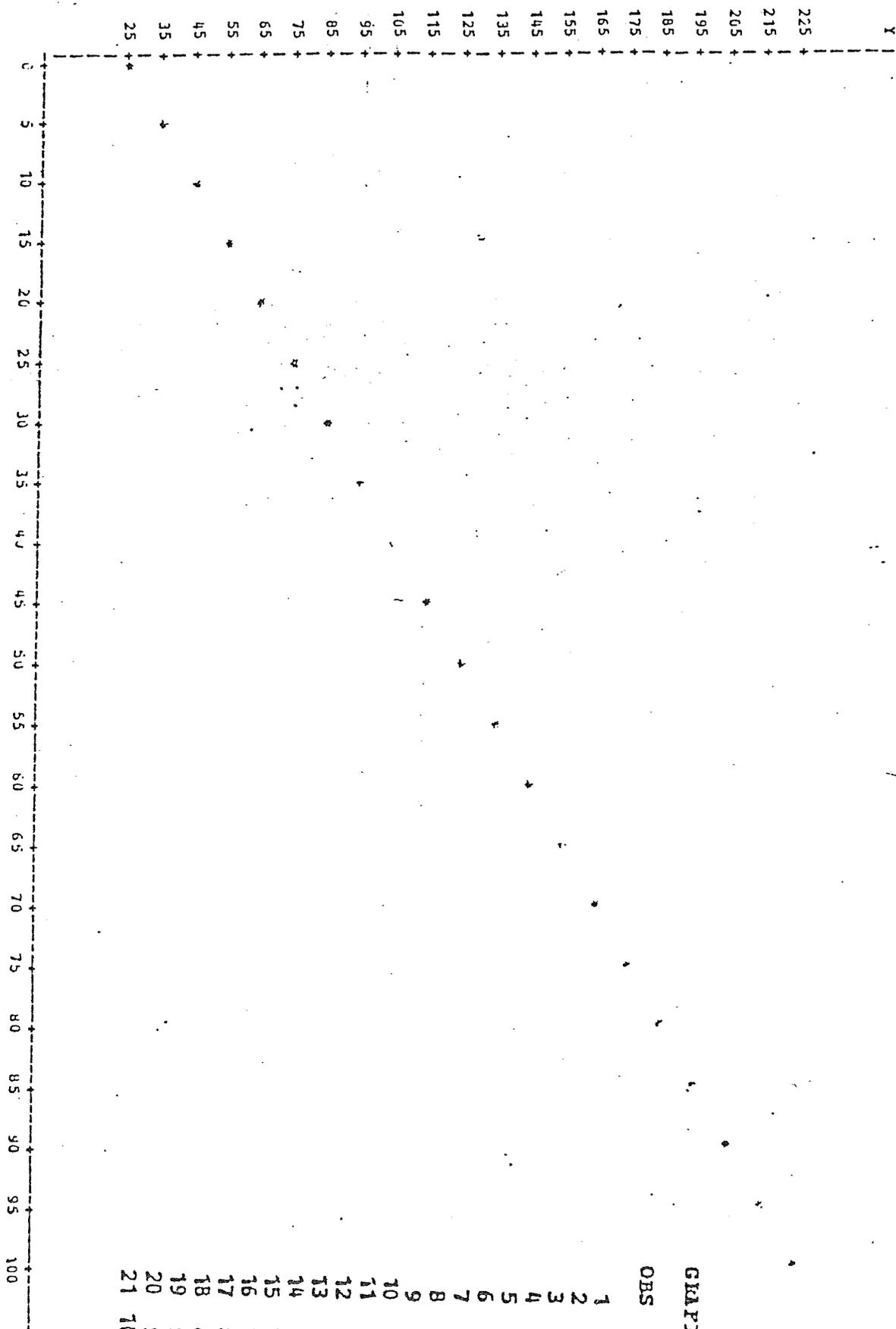


GRAFICO DI $Y=2X+25$

OBS	X	Y
1	0	25
2	5	35
3	10	45
4	15	55
5	20	65
6	25	75
7	30	85
8	35	95
9	40	105
10	45	115
11	50	125
12	55	135
13	60	145
14	65	155
15	70	165
16	75	175
17	80	185
18	85	195
19	90	205
20	95	215
21	100	225

Alcuni cenni alle funzioni SAS

Il SAS mette a disposizione una serie di funzioni (functions) che possono essere utilizzate nelle espressioni numeriche e non, come operatori speciali. Una funzione opera su argomenti e fornisce in uscita un valore.

Le funzioni che operano su argomenti esclusivamente numerici si dicono di tipo numerico, viceversa di tipo carattere. Esistono anche funzioni che elaborano date, tempi o date e tempi; esistono funzioni trigonometriche, probabilistiche, statistiche, ecc.

Sintassi:

```
NOME-FUNZIONE( argom1, argom2, ..., argomN );  
NOME-FUNZIONE( OF variab1-variabN );  
NOME-FUNZIONE( OF variab variab variab );
```

Esempi di funzioni numeriche:

```
Y=ABS(X);  
TOTSTIP=SUM(OF STIPEN1-STIPEN12);  
MINORE=MIN(X Y Z);  
RADQUAD=SQRT(NUMERO);  
MEDIA=MEAN(OF X1-X160);
```

Esempi di funzioni carattere:

```
CODICE='A360';  
PNUMER=SUBSTR(CODICE,1,1);
```

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
  BY SESSO;  
PROC PRINT DATA=SORT.CLASSE3;
```

OBS	NO ME	SESSO	ETA	SCRITTO	ORALE
1	TOZZINI LOREDANA	F	14	5	6
2	BERTINI MARIA	F	14	6	6
3	GHELARDI ANNA	F	14	6	7
4	BIGOTTI MIRIANA	F	14	5	6
5	FANTONI CINZIA	F	13	6	6
6	BALLONI FRANCA	F	14	7	8
7	VALLINI CLAUDIA	F	14	5	5
8	BELLANDI ANGELA	F	14	6	6
9	MORI GINO	M	14	5	5
10	SORRI DANILO	M	13	7	6
11	LIPERINI FABRIZIO	M	15	-	5
12	GAEZELLI PAOLO	M	14	-	-
13	ANIA GIUSEPPE	M	15	4	4
14	MICHELETTI BRUNO	M	14	6	5
15	GARZELLI VITTORIO	M	14	6	6
16	TONINI RENZO	M	16	4	5
17	ACERBI NILO	M	14	6	5
18	GINI ALDO	M	15	6	6
19	CORONA GIOVANNI	M	14	6	5
20	MATTEI DINO	M	14	5	5
21	MARRAZZO UMBERTO	M	14	5	6

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
  BY SESSO NOME;  
PROC PRINT DATA=SORT.CLASSE3;
```

OBS	NOME	SESSO	ETA	SCRITTO	ORALE
1	BALLONI FRANCA	F	14	7	8
2	BELLANDI ANGELA	F	14	6	6
3	BERTINI MARIA	F	14	6	6
4	BIGOTTI MIRIANA	F	14	5	6
5	FANTONI CINZIA	F	13	6	6
6	GHELARDI ANNA	F	14	6	7
7	TOZZINI LOREDANA	F	14	5	6
8	VALLINI CLAUDIA	F	14	5	5
9	ACERBI NILO	M	14	6	5
10	ANIA GIUSEPPE	M	15	4	4
11	CORONA GIOVANNI	M	14	6	5
12	GARZELLI PAOLO	M	14	.	.
13	GARZELLI VITTORIO	M	14	6	6
14	GINI ALDO	M	15	6	6
15	LIPERINI FABRIZIO	M	15	.	5
16	MARRAZZO UMBERTO	M	14	5	6
17	MATTEI DINO	M	14	5	5
18	MICHELETTI BRUNO	M	14	6	5
19	MORI GINO	M	14	5	5
20	SOERI DANILO	M	13	7	6
21	TONINI RENZO	M	16	4	5

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
  BY SESSO DESCENDING ETA HOME;  
PROC PRINT DATA=SORTE.CLASSE3;
```

OBS	NOME	SESSO	ETA	SCRITTO	ORALE
1	BALLONI FRANCA	F	14	7	8
2	BELLANDI ANGELA	F	14	6	6
3	BERTINI MARIA	F	14	6	6
4	BIGOTTI MIRIANA	F	14	5	6
5	GHELARDI ANNA	F	14	6	7
6	TOZZINI LOREDANA	F	14	5	6
7	VALLINI CLAUDIA	F	14	5	5
8	FANTONI CINZIA	F	13	6	6
9	TONINI RENZO	M	16	4	5
10	ANIA GIUSEPPE	M	15	4	4
11	GINI ALDO	M	15	6	6
12	LIPERINI FABRIZIO	M	15	.	5
13	ACERBI NILO	M	14	6	5
14	CORONA GIOVANNI	M	14	6	5
15	GARZELLI PAOLO	M	14	.	.
16	GARZELLI VITTORIO	M	14	6	6
17	MARRAZZO UMBERTO	M	14	5	6
18	MATTEI DIINO	M	14	5	5
19	MICHELETTI BRUNO	M	14	6	5
20	MORI GINO	M	14	5	5
21	SORRI DANILO	M	13	7	6

procedura print

```
PROC PRINT DATA=MEDIA.CLASSE3 N;  
TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';  
VAR NOME ETA;
```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

OBS	NOME	ETA
1	MORI GINO	14
2	SORRI DANILO	13
3	LIPERINI FABRIZIO	15
4	TOZZINI LOREDANA	14
5	BERTINI MARIA	14
6	GARZELLI PAOLO	14
7	GHELARDI ANNA	14
8	ANIA GIUSEPPE	15
9	KICHELETTI BRUNO	14
10	GARZELLI VITTORIO	14
11	BIGOTTI MIRIANA	14
12	TONINI RENZO	16
13	FANTONE CINZIA	13
14	BALLONI FRANCA	14
15	ACERBI NILO	14
16	GINI ALDO	15
17	VALLINI CLAUDIA	14
18	CORONA GIOVANNI	14
19	MATTEI DINO	14
20	BELLANDI ANGELA	14
21	MARRAZZO UMBERTO	14

N=21

procedura print

```
PROC PRINT DATA=MEDIA,CLASSE3 DOUBLE;  
TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';  
VAR ETA SCRITTO;  
ID NOME;
```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

NOME	ETA	SCRITTO
MORI GINO	14	5
SORRI DANILO	13	7
LIPERINI FABRIZIO	15	.
TOZZINI LOREDANA	14	5
BERTINI MARIA	14	6
GARZELLI PAOLO	14	.
GHELARDI ANNA	14	6
ANIA GIUSEPPE	15	4
MICHELETTI BRUNO	14	6
GARZELLI VITTORIO	14	6
BIGOTTI MIRIANA	14	5
TONINI RENZO	16	4
FANTONI CINZIA	13	6
BALLONI FRANCA	14	7
ACERBI NILO	14	6
GINI ALDO	15	6
VALLINI CLAUDIA	14	5
CORONA GIOVANNI	14	6
HATTEI DINO	14	5
BELLANDI ANGELA	14	6
MARRAZZO UMBERTO	14	5

procedura print

```
PROC PRINT DATA=MEDIA.CLASSE3 SPLIT=/:  
TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';  
LABEL NOME = 'COGNOME E NOME';  
SCRITTO = 'VOTO SCRITTO DI ITALIANO';  
ID NOME;
```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

COGNOME E NOME	SESSO	ETA	VOTO SCRITTO DI ITALIANO	ORALE
----------------	-------	-----	-----------------------------	-------

MORI GINO	M	14	5	5
SORRI DANILO	M	13	7	6
LIPERINI FABRIZIO	M	15	.	5
TOZZINI LOREDANA	F	14	5	6
BERTINI MARIA	F	14	6	6
GARZELLI PAOLO	M	14	.	.
GHELARDI ANNA	F	14	6	7
ANIA GIUSEPPE	M	15	4	4
MICHELETTI BRUNO	M	14	6	5
GARZELLI VITTORIO	M	14	6	6
BIGOTTI MIRIANA	F	14	5	6
TONINI RENZO	M	16	4	5
FANTONI CINZIA	F	13	6	6
BALLONI FRANCA	F	14	7	8
ACERBI NILO	M	14	6	5
GINI ALDO	M	15	6	6
VALLINI CLAUDIA	F	14	5	5
CORONA GIOVANNI	M	14	6	5
MATTEI DINO	M	14	5	5
BELLANDI ANGELA	F	14	6	6
MARRAZZO UMBERTO	M	14	5	6

procedura print

```

PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;
  BY SESSO;
PROC PRINT DATA=SORT.CLASSE3 UNIFORM SPLIT=7;
  TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';
  LABEL NOBE = 'COGNOME E NOME';
       SCRITTO = 'VOTO SCRITTO/DI ITALIANO';
  ID NOBE;
  BY SESSO;
  
```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

SESSO=F

COGNOME E NOME	ETA	VOTO SCRITTO DI ITALIANO	ORALE
BALLONI FRANCA	14	7	8
BELLANDI ANGELA	14	6	6
BERTINI MARIA	14	6	6
BIGOTTI MIRIANA	14	5	6
GHELARDI ANNA	14	6	7
TOZZINI LOREDANA	14	5	6
VALLINI CLAUDIA	14	5	5
FANTONI CINZIA	13	6	6

SESSO=M

COGNOME E NOME	ETA	VOTO SCRITTO DI ITALIANO	ORALE
TONINI RENZO	16	4	5
ANIA GIUSEPPE	15	4	4
GINI ALDO	15	6	6
LIPERINI FABRIZIO	15	.	5
ACERBI NILO	14	6	5
CORONA GIOVANNI	14	6	5
GARZELLI PAOLO	14	.	.
GARZELLI VITTORIO	14	6	6
MARRAZZO UMBERTO	14	5	6
MATTEI DINO	14	5	5
MICHELETTI BRUNO	14	6	5
MORI GINO	14	5	5
SORRI DANILO	13	7	6

procedura print

```

PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;
  BY SESSO;
PROC PRINT DATA=SORT.CLASSE3 SPLIT=/:
  TITLE 'SOMMA DEI VOTI DI ITALIANO DELLA CLASSE III A';
  LABEL NONE = 'COGNOME E NOME'
        ETA = 'ETA'
        SCRITTO = 'VOTO SCRITTO/DI ITALIANO'
        ORALE = 'VOTO ORALE/DI ITALIANO';
  ID NONE;
  BY SESSO;
  SUM SCRITTO ORALE;

```

SOMMA DEI VOTI DI ITALIANO DELLA CLASSE III A

SESSO=F

COGNOME E NOME	ETA'	VOTO SCRITTO DI ITALIANO	VOTO ORALE DI ITALIANO
BALLONI FRANCA	14	7	8
BELLANDI ANGELA	14	6	6
BERTINI MARIA	14	6	6
BIGOTTI MIRIANA	14	5	6
GHELARDI ANNA	14	6	7
TOZZINI LOREDANA	14	5	6
VALLINI CLAUDIA	14	5	5
FANTONI CINZIA	13	6	6
-----		46	50
SESSO			

SESSO=M

COGNOME E NOME	ETA'	VOTO SCRITTO DI ITALIANO	VOTO ORALE DI ITALIANO
TONINI RENZO	16	4	5
ANIA GIUSEPPE	15	4	4
GINI ALDO	15	6	6
LIPERINI FABRIZIO	15	-	5
ACERBI NILO	14	6	5
CORONA GIOVANNI	14	6	5
GARZELLI PAOLO	14	-	-
GARZELLI VITTORIO	14	6	6
MARRAZZO UMBERTO	14	5	6
MATTEI DINO	14	5	5
MICHELETTI BRUNO	14	6	5
MORI GINO	14	5	5
SORRI DANILO	13	7	6
-----		60	63
SESSO		===	===
		106	113

procedura format

```

PROC FORMAT;
  VALUE $SEX
    F = 'FEMMINA';
    M = 'MASCIO';
    OTHER = 'NEUTRO';
  VALUE VOTO
    . = 'NON CLASSIFICATO';
    0-5 = 'INSUFFICIENTE';
    6,7 = 'SUFFICIENTE';
    8-HIGH = 'BUONO';
  PICTURE ANNI
    LOW-HIGH = '0999';
PROC PRINT DATA=MEDIA.CLASSE3 SPLIT=7;
  TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';
  LABEL COGNOME = 'COGNOME E NOME';
  ETA = 'ETA';
  SCRITTO = 'VOTO SCRITTO/DI ITALIANO';
  ORALE = 'VOTO ORALE/DI ITALIANO';
  FORMAT $SEX.;
  FORMAT SCRITTO VOTO.;
  FORMAT ORALE VOTO.;
  FORMAT ANNI.;
  ID NOME;

```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

COGNOME E NOME	SESSO	ETA'	VOTO SCRITTO DI ITALIANO	VOTO ORALE DI ITALIANO
BALLONI FRANCA	FEMMINA	014	SUFFICIENTE	BUONO
BELLANDI ANGELA	FEMMINA	014	SUFFICIENTE	SUFFICIENTE
BERTINI MARIA	FEMMINA	014	SUFFICIENTE	SUFFICIENTE
BIGOTTI MIRIANA	FEMMINA	014	INSUFFICIENTE	SUFFICIENTE
GHELARDI ANNA	FEMMINA	014	SUFFICIENTE	SUFFICIENTE
TOZZINI LOREDANA	FEMMINA	014	INSUFFICIENTE	SUFFICIENTE
VALLINI CLAUDIA	FEMMINA	014	INSUFFICIENTE	INSUFFICIENTE
FANTONI CINZIA	FEMMINA	013	SUFFICIENTE	SUFFICIENTE
TONINI RENZO	MASCIO	016	INSUFFICIENTE	INSUFFICIENTE
ANIA GIUSEPPE	MASCIO	015	INSUFFICIENTE	INSUFFICIENTE
GINI ALDO	MASCIO	015	SUFFICIENTE	SUFFICIENTE
LIPERINI FABRIZIO	MASCIO	015	NON CLASSIFICATO	INSUFFICIENTE
ACERBI NILO	MASCIO	014	SUFFICIENTE	INSUFFICIENTE
CORONA GIOVANNI	MASCIO	014	SUFFICIENTE	INSUFFICIENTE
GARZELLI PAOLO	MASCIO	014	NON CLASSIFICATO	NON CLASSIFICATO
GARZELLI VITTORIO	MASCIO	014	SUFFICIENTE	SUFFICIENTE
MARRAZZO UMBERTO	MASCIO	014	INSUFFICIENTE	SUFFICIENTE
MATTEI DINO	MASCIO	014	INSUFFICIENTE	INSUFFICIENTE
MICHELETTI BRUNO	MASCIO	014	SUFFICIENTE	INSUFFICIENTE
MORI GINO	MASCIO	014	INSUFFICIENTE	INSUFFICIENTE
SORRI DANILO	MASCIO	013	SUFFICIENTE	SUFFICIENTE

procedura means

PROC MEANS DATA=MEDIA CLASSE3;

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD. ERROR OF MEAN
ETA	21	14.142857	0.654653	13.000000	16.000000	0.142857
SCRITTO	19	5.578947	0.837707	4.000000	7.000000	0.192183
ORALE	20	5.650000	0.875093	4.000000	8.000000	0.195676

procedura means

```
PROC MEANS DATA=MEDIA CLASSES MAXDEC=2;  
VAR SCRITTO ORALE;
```

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN
SCRITTO	19	5.58	0.84	4.00	7.00	0.19
ORALE	20	5.65	0.88	4.00	8.00	0.20

procedura means

```
PROC MEANS DATA=MEDIA.CLASSE3 MAXDEC=2 N SUM RANGE MEAN VAR NMISS;  
TITLE 'MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A';  
VAR SCRITTO ORALE;
```

MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A

VARIABLE	N	SUM	RANGE	MEAN	VARIANCE	N MISSING
SCRITTO	19	106.00	3.00	5.58	0.70	2
ORALE	20	113.00	4.00	5.65	0.77	1

procedura means

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
  BY SESSO;  
PROC MEANS DATA=SORT.CLASSE3 MAXDEC=2;  
  TITLE *MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A*;  
  VAR SCRITTO ORALE;  
  BY SESSO;
```

MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN

SESSO=F						
SCRITTO	8	5.75	0.71	5.00	7.00	0.25
ORALE	8	6.25	0.89	5.00	8.00	0.31

SESSO=M						
SCRITTO	11	5.45	0.93	4.00	7.00	0.28
ORALE	12	5.25	0.62	4.00	6.00	0.18

procedura freq

PROC FREQ DATA=MEDIA.CLASSE3:

NOME	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
ACERBI NILO	1	1	4.762	4.762
ANIA GIUSEPPE	1	2	4.762	9.524
BALLONI FRANCA	1	3	4.762	14.286
BELLANDI ANGELA	1	4	4.762	19.048
BERTINI MARIA	1	5	4.762	23.810
BIGOTTI MIRIANA	1	6	4.762	28.571
CORONA GIOVANNI	1	7	4.762	33.333
FANTONI CINZIA	1	8	4.762	38.095
GARZELLI PAOLO	1	9	4.762	42.857
GARZELLI VITTORI	1	10	4.762	47.619
GHELARDI ANNA	1	11	4.762	52.381
GINI ALDO	1	12	4.762	57.143
LIPERINI FABRIZI	1	13	4.762	61.905
MARRAZZO UMBERTO	1	14	4.762	66.667
MATTEI DINO	1	15	4.762	71.429
MICHELETTI BRUNO	1	16	4.762	76.190
MORI GINO	1	17	4.762	80.952
SORRI DANILO	1	18	4.762	85.714
TONINI RENZO	1	19	4.762	90.476
TOZZINI LOREDANA	1	20	4.762	95.238
VALLINI CLAUDIA	1	21	4.762	100.000

SESSO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
F	8	8	38.095	38.095
M	13	21	61.905	100.000

ETA	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
13	2	2	9.524	9.524
14	15	17	71.429	80.952
15	3	20	14.286	95.238
16	1	21	4.762	100.000

SCRITTO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
-	2	-	-	-
4	2	2	10.526	10.526
5	6	8	31.579	42.105
6	9	17	47.368	89.474
7	2	19	10.526	100.000

ORALE	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
-	1	-	-	-
4	1	1	5.000	5.000
5	8	9	40.000	45.000
6	9	18	45.000	90.000
7	1	19	5.000	95.000
8	1	20	5.000	100.000

procedura freq

```
PROC FREQ DATA=MEDIA.CLASSE3 ORDER=FREQ;  
TITLE "FREQUENZA DELLE VARIABILI";  
TABLE SESSO ETA;
```

FREQUENZA DELLE VARIABILI

SESSO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
M	13	13	61.905	61.905
F	8	21	38.095	100.000

ETA	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
14	15	15	71.429	71.429
15	3	18	14.286	85.714
13	2	20	9.524	95.238
16	1	21	4.762	100.000

procedura freq

```

PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;
  BY SESSO;
PROC FREQ DATA=SORT.CLASSE3 ORDER=FREQ;
  TITLE 'FREQUENZA DELLE VARIABILI';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO'
        ORALE = 'VOTO ORALE DI ITALIANO';
  TABLE SCRITTO;
  BY SESSO;
  
```

FREQUENZA DELLE VARIABILI SESSO=F

VOTO SCRITTO DI ITALIANO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
6	4	4	50.000	50.000
5	3	7	37.500	87.500
7	1	8	12.500	100.000

FREQUENZA DELLE VARIABILI SESSO=M

VOTO SCRITTO DI ITALIANO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
.	2	2	45.455	45.455
6	5	7	27.273	72.727
5	3	10	18.182	90.909
4	2	12	9.091	100.000
7	1	13		

procedura freq

```

PROC FREQ DATA=MEDIA.CLASSE3 ORDER=FREQ;
  TITLE 'FREQUENZA DELLE VARIABILI';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO'
        ORALE = 'VOTO ORALE DI ITALIANO';
  TABLE SESSO*SCRITTO;
  
```

FREQUENZA DELLE VARIABILI

TABLE OF SESSO BY SCRITTO

SESSO SCRITTO VOTO SCRITTO DI ITALIANO

FREQUENCY						TOTAL
PERCENT						
ROW PCT						
COL PCT	-	6	5	7	4	
M	2	5	3	1	2	11
	-	26.32	15.79	5.26	10.53	57.89
	-	45.45	27.27	9.09	18.18	
	-	55.56	50.00	50.00	100.00	
F	0	4	3	1	0	8
	-	21.05	15.79	5.26	0.00	42.11
	-	50.00	37.50	12.50	0.00	
	-	44.44	50.00	50.00	0.00	
TOTAL	-	9	6	2	2	19
	-	47.37	31.58	10.53	10.53	100.00

procedura freq

```

PROC FREQ DATA=MEDIA.CLASSE3 ORDER=FREQ;
  TITLE 'FREQUENZA DELLE VARIABILI';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO'
        ORALE = 'VOTO ORALE DI ITALIANO';
  TABLE SESSO*(SCRITTO ORALE)/NOFREQ NOROW NOCOL;
  
```

FREQUENZA DELLE VARIABILI

TABLE OF SESSO BY SCRITTO

SESSO	SCRITTO	VOTO SCRITTO DI ITALIANO				TOTAL
PERCENT		6	5	7	4	
M		26.32	15.79	5.26	10.53	57.89
F		21.05	15.79	5.26	0.00	42.11
TOTAL		9	6	2	2	19
		47.37	31.58	10.53	10.53	100.00

FREQUENZA DELLE VARIABILI

TABLE OF SESSO BY ORALE

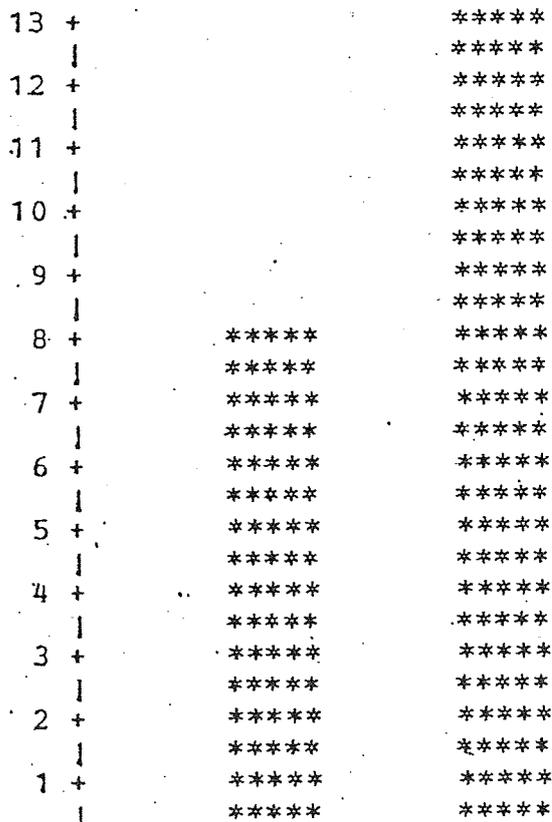
SESSO	ORALE	VOTO ORALE DI ITALIANO				TOTAL	
PERCENT		6	5	8	7	4	
M		20.00	35.00	0.00	0.00	5.00	60.00
F		25.00	5.00	5.00	5.00	0.00	40.00
TOTAL		9	8	1	1	1	20
		45.00	40.00	5.00	5.00	5.00	100.00

procedura chart

```
PROC CHART DATA=MEDIA.CLASSE3;  
VBAR SESSO / TYPE=SUM;
```

FREQUENCY BAR CHART

FREQUENCY



F M

SESSO

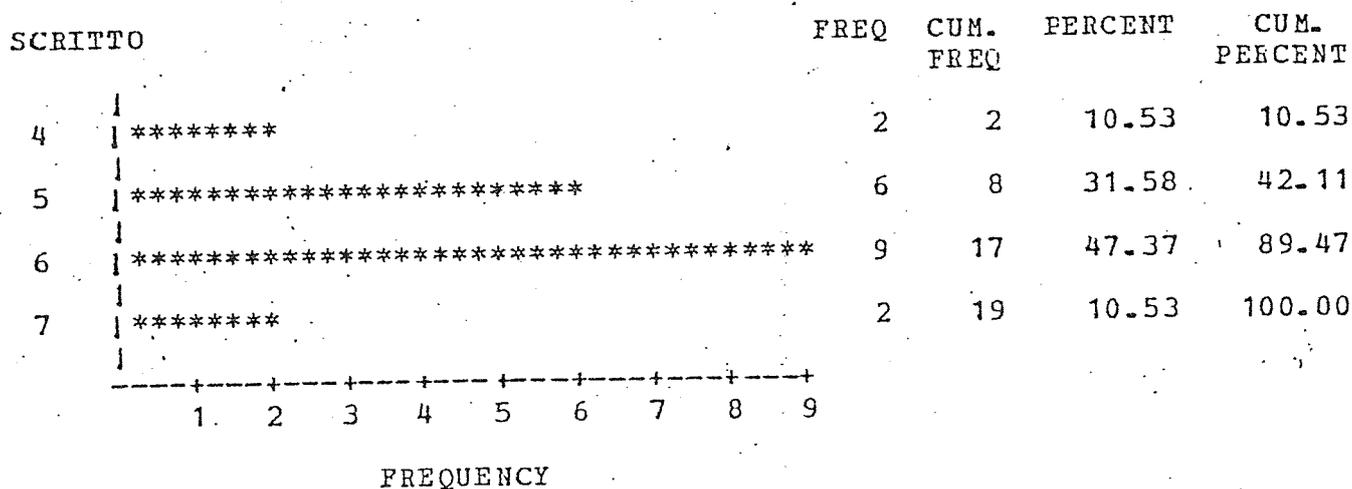
procedura chart

```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';
  ORALE = 'VOTO ORALE DI ITALIANO';
  EBAN SCRITTO ORALE / TYPE=FREQ DISCRETE;
  
```

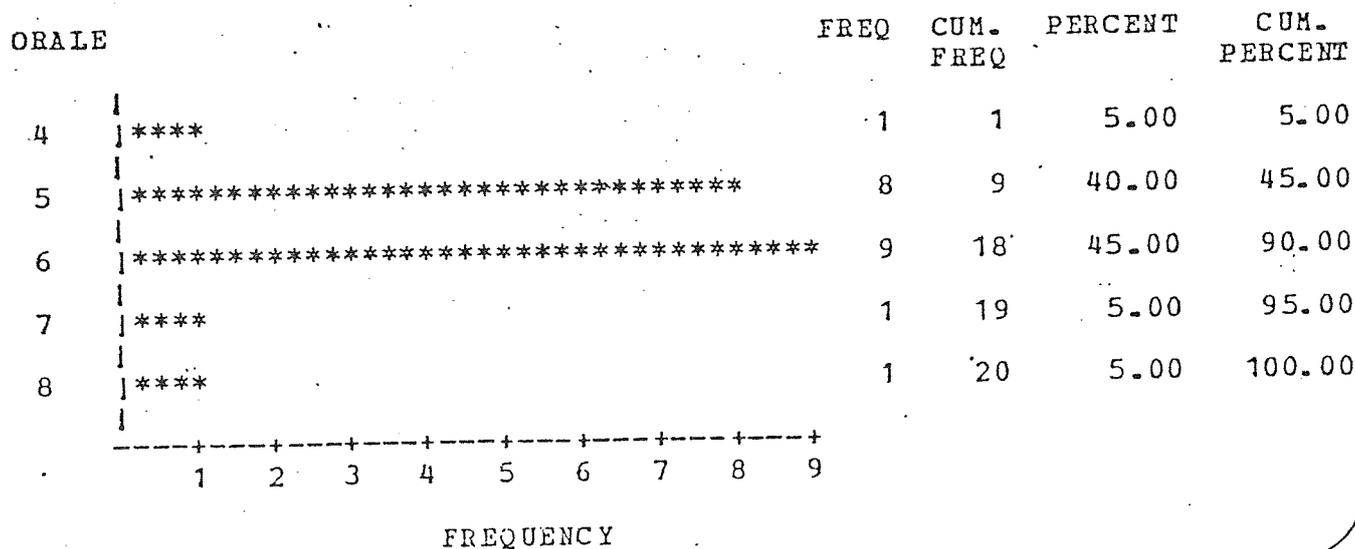
DISTRIBUZIONE DEI VOTI DI ITALIANO

FREQUENCY BAR CHART



DISTRIBUZIONE DEI VOTI DI ITALIANO

FREQUENCY BAR CHART



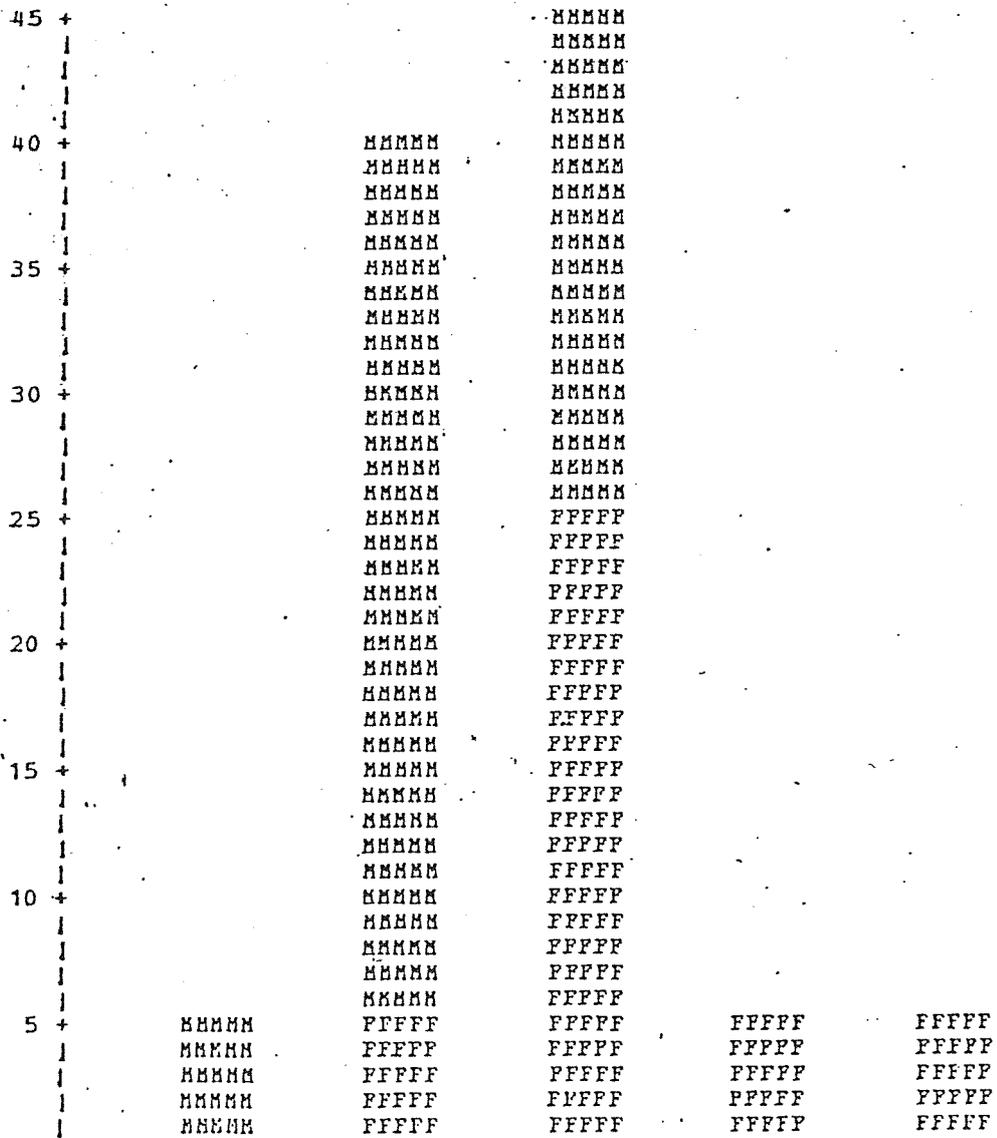
procedura chart

```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL ORALE = 'VOTO ORALE DI ITALIANO';
  VBAR ORALE / TYPE=PCT DISCRETE SUBGROUP=SESSO;
  
```

PERCENTAGE BAR CHART

PERCENTAGE



4 5 6 7 8

ORALE

procedura chart

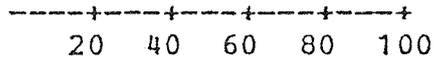
```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';
  HBAR SCRITTO / TYPE=CPCT DISCRETE GROUP=SESSO;
  
```

DISTRIBUZIONE DEI VOTI DI ITALIANO

CUMULATIVE PERCENTAGE BAR CHART

SESSO	SCRITTO	FREQ	CUM. FREQ	PERCENT	CUM. PERCENT
F	4	0	0	0.00	0.00
	5	3	3	15.79	15.79
	6	4	7	21.05	36.84
	7	1	8	5.26	42.11
M	4	2	10	10.53	52.63
	5	3	13	15.79	68.42
	6	5	18	26.32	94.74
	7	1	19	5.26	100.00



CUMULATIVE PERCENTAGE

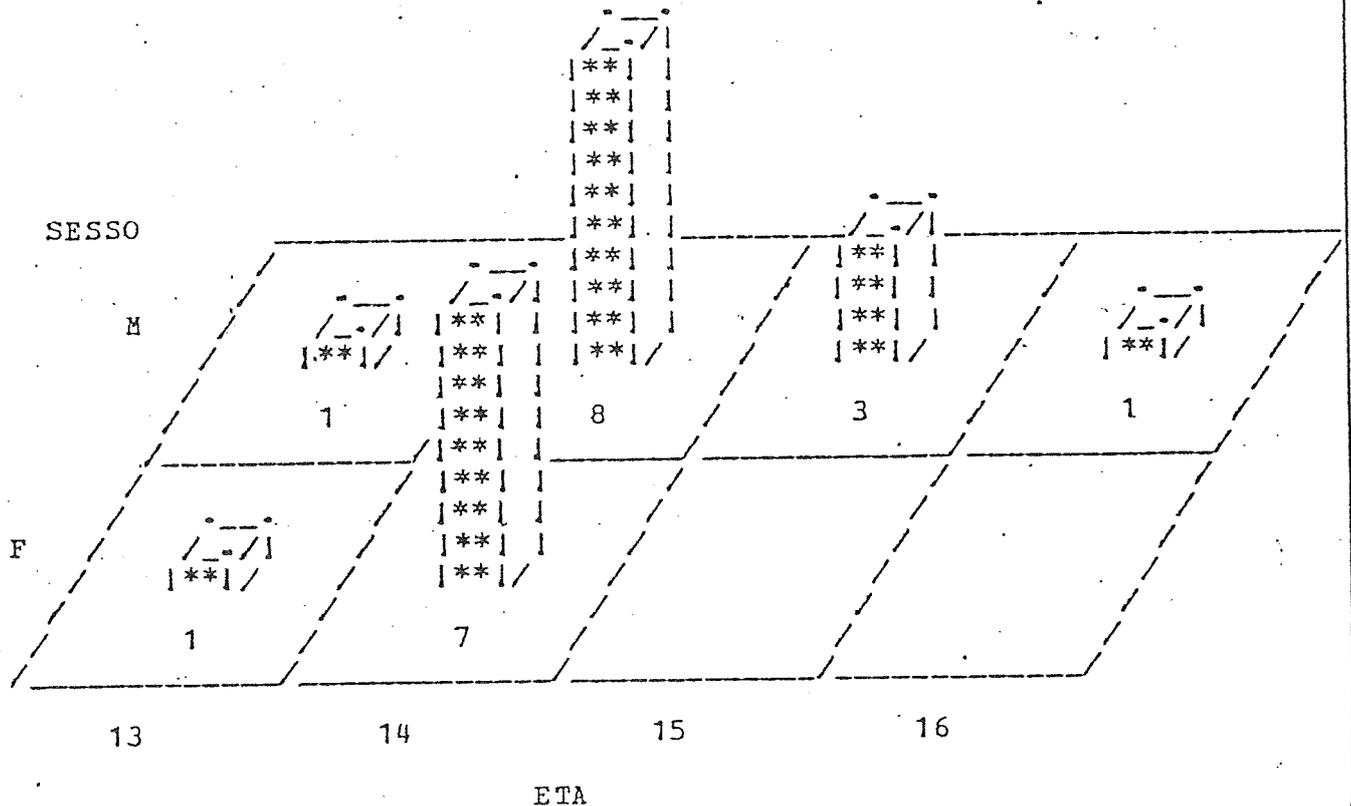
procedura chart

```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DELLE ETA';
  LABEL ETA = 'ETA';
  BLOCK ETA / TYPE=FRFQ DISCRETE GROUP=SESSO;
  
```

DISTRIBUZIONE DELLE ETA'

FREQUENCY BLOCK CHART



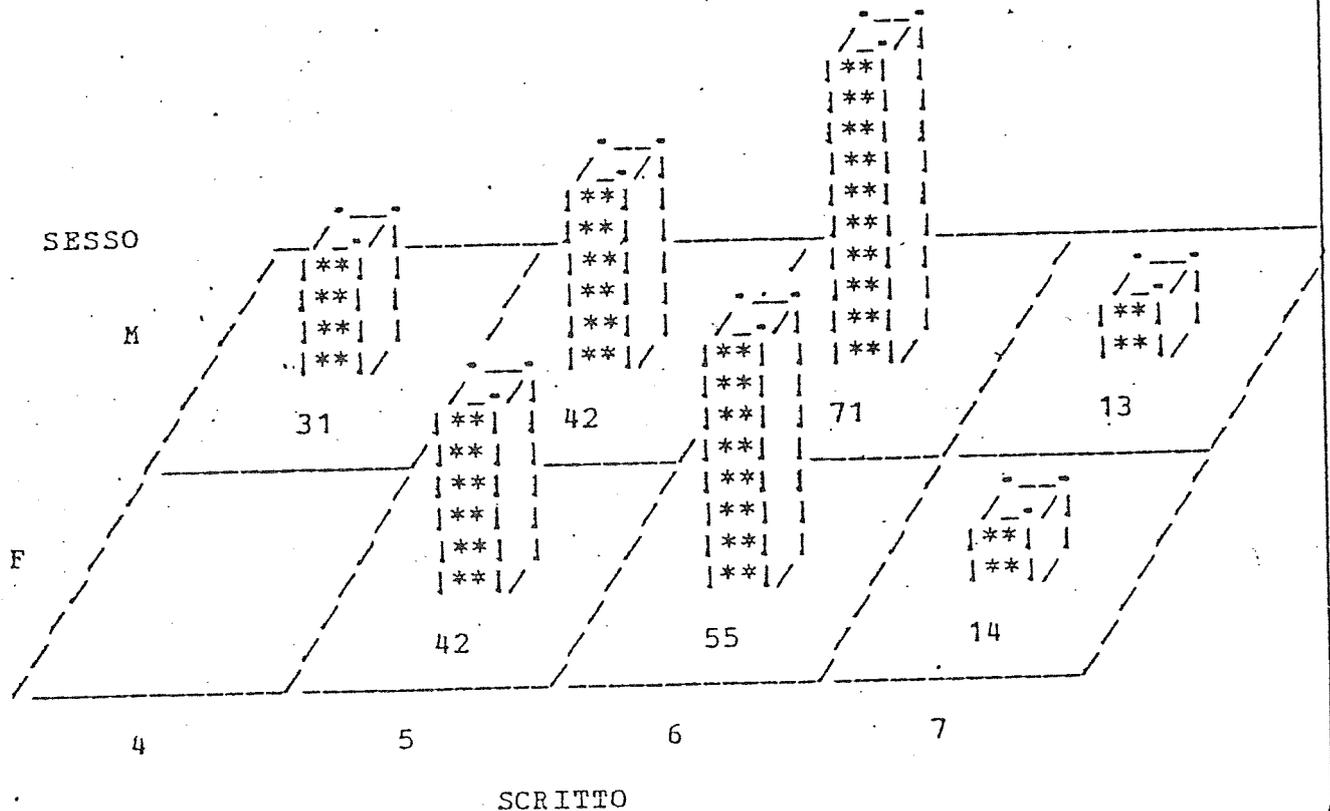
procedura chart

```

PROC CHART DATA=MEDIA_CLASSES;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';
  BLOCK SCRITTO / TYPE=Freq DISCRETE GROUP=SESSO SUMVAR=ETA;
  
```

DISTRIBUZIONE DEI VOTI DI ITALIANO

BLOCK CHART OF ETA SUMS

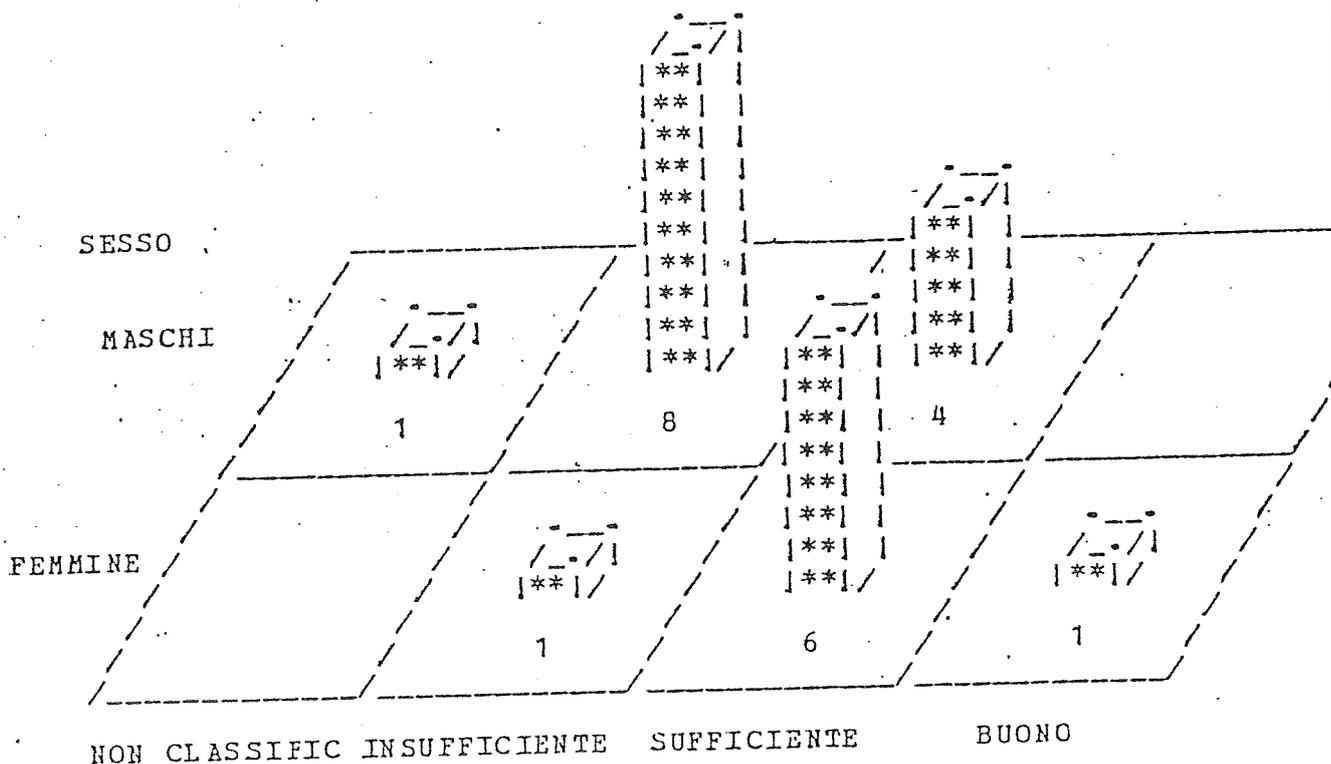


procedura format

```

PROC FORMAT;
  VALUE $SEX
    F = 'FEMMINE'
    M = 'MASCHI'
    OTHER = 'NEUTRO';
  VALUE VOTO
    - = 'NON CLASSIFICATO'
    0-5 = 'INSUFFICIENTE'
    6,7 = 'SUFFICIENTE'
    8-HIGH = 'BUONO';
PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO DELLA CLASSE III A';
  LABEL ORALE = 'VOTO ORALE DI ITALIANO';
  FORMAT SESSO $SEX;
  FORMAT ORALE VOTO;
  BLOCK ORALE / DISCRETE TYPE=FREQ GROUP=SESSO MISSING;
  
```

DISTRIBUZIONE DEI VOTI DI ITALIANO DELLA CLASSE III A
 FREQUENCY BLOCK CHART

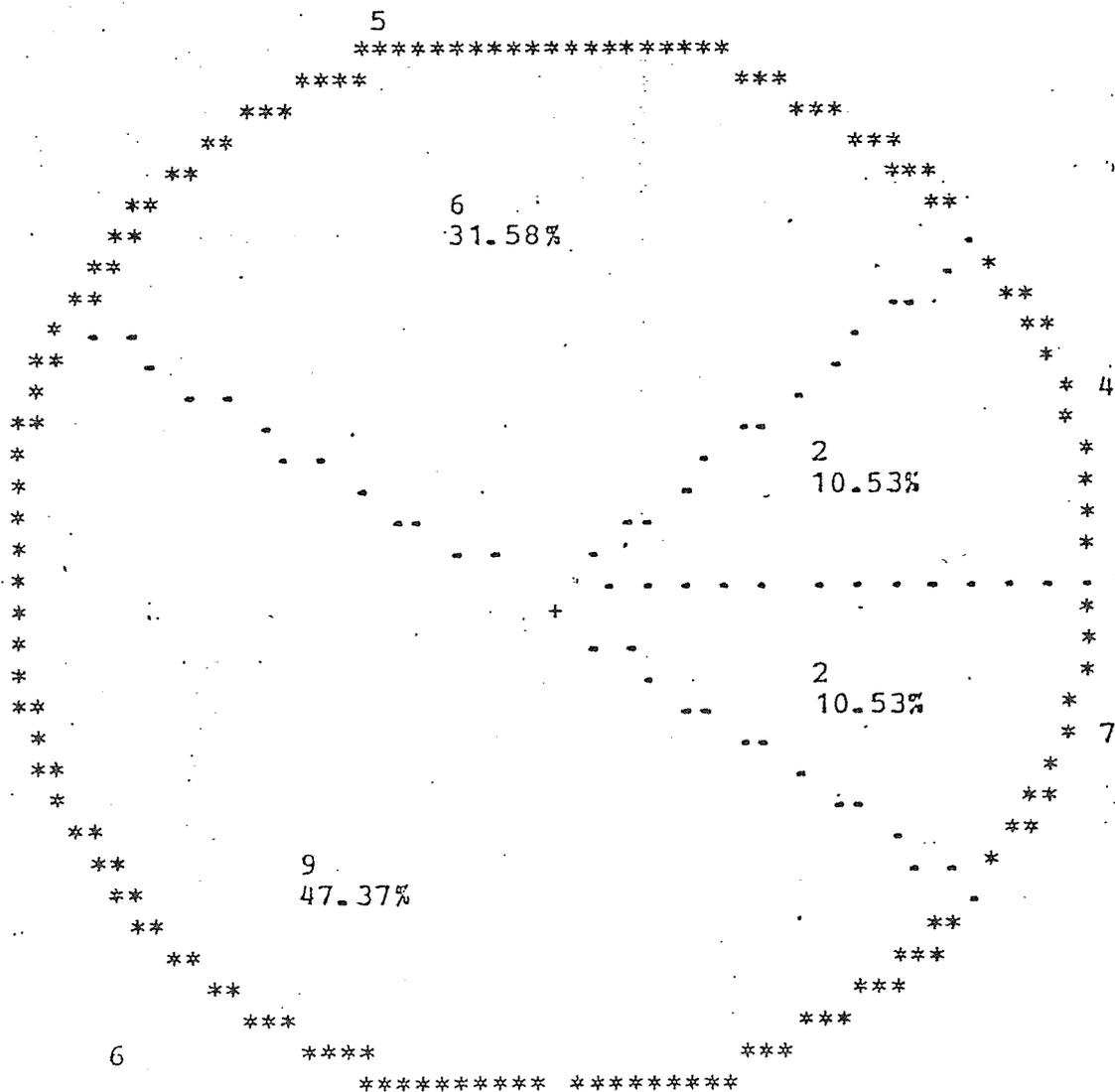


procedura chart

```
PROC CHART DATA=MEDIA.CLASSE3;  
TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';  
LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';  
PIE SCRITTO/ TYPE=REQ DISCRETE;
```

DITRIBUZIONE DEI VOTI DI ITALIANO

FREQ PIE CHART OF SCRITTO

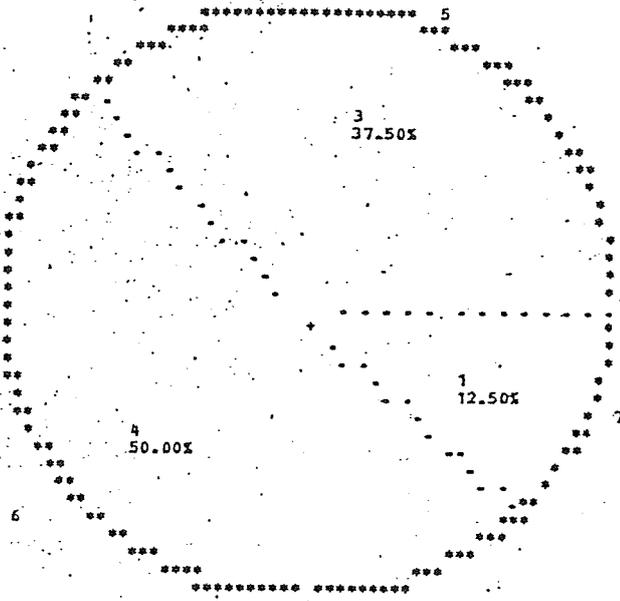


procedura chart

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
BY SSSO;  
PROC CHART DATA=SORT.CLASSE3;  
TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';  
LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';  
PIE SCRITTO/ TYPE=REQ DISCRETE;  
BY SSSO;
```

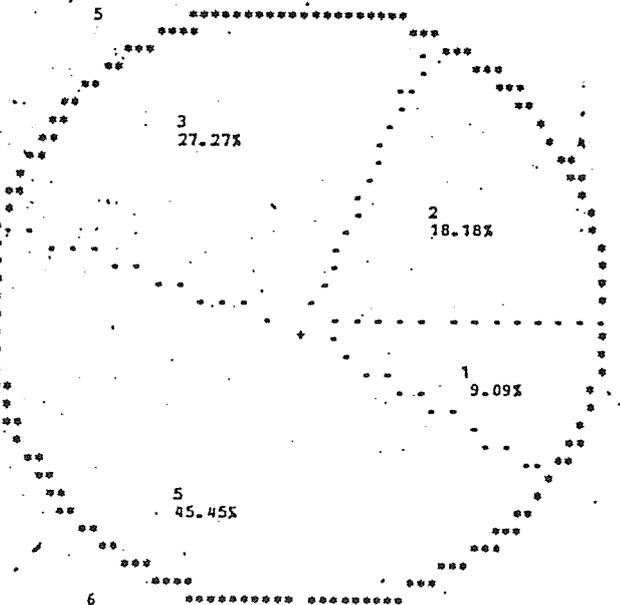
DISTRIBUZIONE DEI VOTI DI ITALIANO
SSSO=F

FREQ PIE CHART OF SCRITTO



DISTRIBUZIONE DEI VOTI DI ITALIANO
SSSO=M

FREQ PIE CHART OF SCRITTO



Dichiarazione SET

La dichiarazione SET permette la lettura di osservazioni da uno o piu' data set di tipo SAS (non si usa la INPUT !!). Nel caso si legga da un solo data set le osservazioni sono lette una per una in maniera sequenziale sino alla fine del data set. Nel caso che i data set siano due o piu' si puo' ottenere la concatenazione o l'interleaving dei data set.

Sintassi:

```
SET nome di data set SAS opzioni;  
    BY lista di variabili;
```

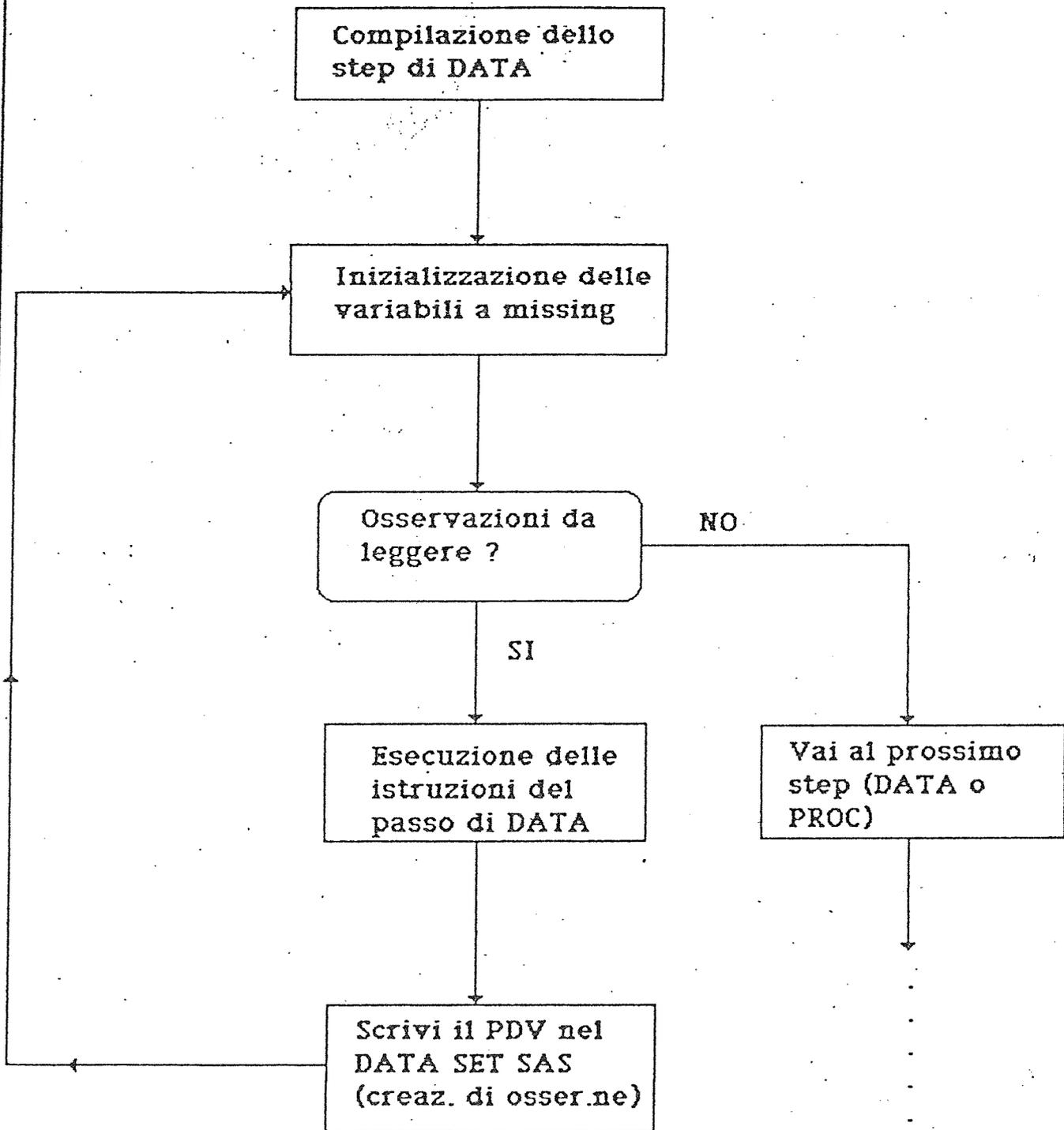
Esempi:

```
SET POPOL;  
SET FATTUR.ANN084;  
SET _LAST_;  
SET DATASET (DROP=NOME STIPEND);
```

```
SET A B; Concatena il data set B al  
data set A
```

```
PROC SORT DATA=MASCHI; BY MATRICOL;  
PROC SORT DATA=FEMMINE; BY MATRICOL;  
DATA TUTTIORD;  
    SET MASCHI FEMMINE;  
    BY MATRICOL;
```

CICLO DI LAVORO DI UN PASSO DI DATA



CAPITOLO 4

PROCEDURE STEP

PROC PRINT options DATA=datasetsAS;

SERVE PER STAMPARE I VALORI DELLE VARIABILI DEL DATASET

L'USO DI OPZIONI E DI STATEMENT DI CONTROLLO PERMETTE LA PERSONALIZZAZIONE DELLE STAMPE;

OPZIONI:

N STAMPA IL NUMERO DI OSSERVAZIONI DEL DATASET

DOUBLE (D) SPAZIATURA DOPPIA TRA UNA RIGA E L'ALTRA

UNIFORM CURA L'UNIFORMITA' DELLA PAGINA DI STAMPA (E' UTILE QUANDO IL DATASET HA MOLTE VARIABILI)

ROUND ARROTONDA I VALORI NUMERICI PRIMA DI TOTALIZZARLI

LABEL LE LABEL ATTRIBUITE ALLE VARIABILI SONO USATE COME INTRESTAZIONE

SPLT= DELIMITATORE INDICANTE IL PUNTO DI DIVISIONE DELLE LABEL

STATEMENTS:

BY LISTA DI VARIABILI;

SUDDIVIDE I QUADRI PER LE VARIABILI INDICATE NELLA LISTA

VAR LISTA DI VARIABILI;

LIMITA LA STAMPA ALLE SOLE VARIABILI DELLA LISTA

ID LISTA DI VARIABILI;

RIMPIAZZA LA NUMERAZIONE PROGRESSIVA CON I VALORI DELLA PRIMA VARIABILE DELLA LISTA

(ANCORA SULLA PROC PRINT)

SUM LISTA DI VARIABILI

TOTALIZZA LE VARIABILI DELLA
LISTA

LO SCOPO PRINCIPALE DELLA STATISTICA CONSISTE NEL COMPIERE UN'INFERENZA CIRCA L'INTERA POPOLAZIONE A PARTIRE DA UN CAMPIONE.

CIASCUN CAMPIONE PUO' ESSERE SEMPLIFICATO DA POCHI NUMERI DESCRITTIVI, CIASCUNO DEI QUALI SI DEFINISCE "STATISTICA CAMPIONARIA".

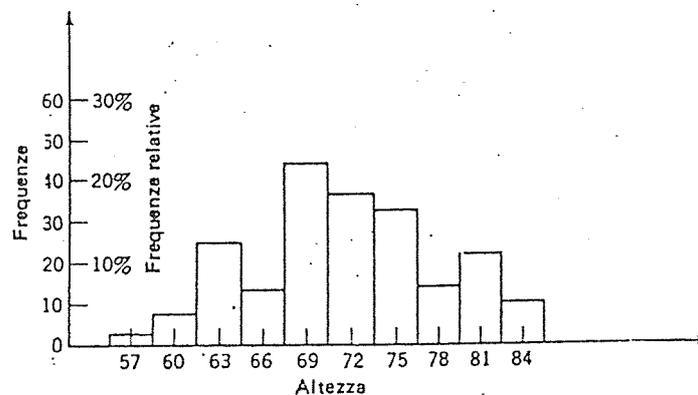
SPESO CONVIENE RAPPRESENTARE I DATI DEL CAMPIONE IN MANIERA SINOTTICA E COMPATTA, IN MODO DA RENDERNE AGEVOLE LO STUDIO; LA TABELLA DI FREQUENZA SOTTOSTANTE NE È UN ESEMPIO:

Frequenze e frequenze relative delle stature di un campione di 200 individui.

N. delle classi	Classi limiti	Classi valore centrale	Frequenze f	Frequenze relative f/n
1	55.5-58.5	57	2	.010
2	58.5-61.5	60	7	.035
3	61.5-64.5	63	22	.110
4		66	13	.065
5		69	44	.220
6		72	36	.180
7		75	32	.160
8		78	13	.065
9		81	21	.105
10	82.5-85.5	84	10	.050

$$\sum f = 200 = n \quad \sum f/n = 1.00$$

UN ALTRO MODO DI RAPPRESENTARE GLI STESSI DATI È QUELLO DI USARE I GRAFICI DI FREQUENZA DEL TIPO MOSTRATO SOTTO:



Distribuzione delle frequenze e delle frequenze relative di un campione di 200 individui.

IL S.A.S. PERMETTE LA REALIZZAZIONE DI TABELLE DI FREQUENZA E DI GRAFICI DI FREQUENZA IN MANIERA SEMPLICE E VELOCE.

TABELLE DI FREQUENZA

LE TABELLE DI FREQUENZA POSSONO ESSERE DI **TRE** TIPI:

1) A 1 DIMENSIONE

2) " 2 DIMENSIONI

3) " 3 " "

PROC FREQ DATA=DATASET SAS;

QUESTA PROCEDURA È USATA PER GENERARE TABELLE DI FREQUENZA AD UNA O PIÙ VIE.

INOLTRE È POSSIBILE OTTENERE:

- DISTRIBUZIONE DEI VALORI DI UNA VARIABILE
- DISTRIBUZIONI CONGIUNTE DI DUE O PIÙ VARIABILI
- FREQUENZE PESATE
- MISURE DI ASSOCIAZIONE E TEST STATISTICI PER TABELLE A DUE VIE
- GENERARE DATASET CON I RISULTATI (INVECE CHE SU STAMPANTE)

STATEMENTS USABILI:

TABLES richieste di tabulazione / opzioni;

ESEMPI: TABELLA AD UNA VIA TABLES SESSO;
 TABELLA A DUE VIE TABLES SESSO*PAGA;
 TABELLA A TRE VIE TABLES CLASSE*SESSO*VOTO;

OPZIONI (ALCUNE):

EXPECTED	NOFREQ	NOCUM	ALL
DEVIATION	NOPERCENT	MISSING	LIST
CELLCHI2	NOROW	NOPRINT	SPARSE
CHISQ	NOCOL	OUT=DATASETSAS	

WEIGHT VARIABILE-PESO;

BY LISTA DI VARIABILI;

PROCEDURA CHART

LA PROCEDURA CHART E' UTILE PER DISEGNARE ISTOGRAMMI
E IDEOGRAMMI (TORTE E STELLE), SULLA STAMPANTE O SUL
TERMINALE;

SI POSSONO OTTENERE LE SEGUENTI FIGURE:

- ISTOGRAMMI ORIZZONTALI
- ISTOGRAMMI VERTICALI
- ISTOGRAMMI A BLOCCHI (TRIDIMENSIONALI)
- TORTE (PIE)
- STELLE

FORMATI

Cos'è un formato? PUO' ESSERE ASSIMILATO AD UN "FILTRO" CHE DEFINISCE "COME" IL VALORE DI UNA VARIABILE DEVE ESSERE PRODOTTO IN USCITA/ENTRATA

ESISTONO FORMATI GIÀ "PRECONFEZIONATI" NEL SAS ED IL LORO SPETTRO DI APPLICAZIONE È MOLTO AMPIO:

ESEMPI: -W. W.D SW. COMMAW.D
 ROMANW. WORDSW. BESTW. ECC. ECC..

L'ASSOCIAZIONE DI UNA FORMATO AD UNA O PIÙ VARIABILI AVVIENE TRAMITE LO STATEMENT FORMAT:

ESEMPIO: FORMAT ALTEZZA PESO 3. ;
 FORMAT NOME S30. ALTEZZA 3. PESO 2.;

N.B.: SE QUESTO STATEMENT VIENE USATO IN UN PASSO DI DATA, ESSO VIENE ASSOCIATO ALLA VARIABILE IN MANIERA PERMANENTE SINO AD UNA NUOVA ASSOCIAZIONE TRAMITE FORMAT. SE VIENE USATA IN UNA PROC, IL SUO EFFETTO SARÀ VALIDO PER LA SOLA DURATA DELLA PROC.

CREAZIONE DI FORMATI "AD HOC"

E' POSSIBILE CREARE UN FORMATO PER I PROPRI SCOPI USANDO LA PROCEDURA:

```
PROC FORMAT opzioni;
```

OPZIONI:

```
PRINT DDNAME=ddname.
```

STATEMENTS USABILI:

```
VALUE nome del formato (opzioni)
```

```
range1=label1
```

```
range2=label2
```

```
.....  
.....;
```

```
PICTURE nome del formato (opzioni)
```

```
range1='picture1' (opzioni)
```

```
range2='picture2' (opzioni)
```

```
.....  
.....;
```

NOTE: INFILE COMPON IS FILE COMPON FILE A1
NOTE: 7 LINES WERE READ FROM INFILE COMPON.
THE MINIMUM LINE LENGTH IS 120.
THE MAXIMUM LINE LENGTH IS 120.

NOTE: DATA SET WORK,GIUNTA HAS 7 OBSERVATIONS AND 20 VARIABLES.
NOTE: THE DATA STATEMENT USED 0.12 SECONDS AND 832K.

```

28      PROC FORMAT;
29      VALUE LISTFORM 01 = PCI
30      02 = PSI
31      03 = PDUP
32      04 = DP
33      05 = 'IND-SINISTRA'
34      06 = PSDI
35      07 = PRI
36      08 = DC
37      09 = PLI
38      10 = 'IND-CENTRO'
39      11 = 'MSI-DN'
40      12 = 'ALTRI-LISTA LOCALE'
41      13 = 'PART. RADICALE';

```

NOTE: THE PROCEDURE FORMAT USED 0.10 SECONDS AND 832K.

CAPITOLO 5

ESEMPI DI PROGRAMMI

data step

DATA MEDIA.CLASSE3;
INPUT NOME \$25.
 SESSO \$ 27-27
 +2 ETA 2.
 @35 SCRITTO 35.
 @40 ORALE 1.;

CARDS;

KORI GINO	M	14	5	5
SORRI DANILO	M	13	7	6
LIPERINI FABRIZIO	M	15	.	5
TOZZINI LOREDANA	F	14	5	6
BERTINI MARIA	F	14	6	6
GARZELLI PAOLO	M	14	.	.
GHELARDI ANNA	F	14	6	7
ANIA GIUSEPPE	M	15	4	4
MICHELETTI BRUNO	M	14	6	5
GARZELLI VITTORIO	M	14	6	6
BIGOTTI MIRIANA	F	14	5	6
TONINI RENZO	M	16	4	5
FANTONI CINZIA	F	13	6	6
BALLONI FRANCA	F	14	7	8
ACERBI NILO	M	14	6	5
GINI ALDO	M	15	6	6.
VALLINI CLAUDIA	F	14	5	5
CORONA GIOVANNI	M	14	6	5
MATTEI DINO	M	14	5	5
BELLANDI ANGELA	F	14	6	6
MARRAZZO UMBERTO	M	14	5	6

;

procedura contents

PROC CONTENTS DATA=MEDIA.CLASSE3;

CONTENTS OF SAS DATA SET MEDIA.CLASSE3

OBSERVATIONS=21 CREATED BY CMS USERID MESURE ON CPUID FF-3081-001964

AT 15:00 TUESDAY, APRIL 30, 1985 BY SAS RELEASE 82.3

FILE=MEDIA CLASSE3 BLKSIZE=8158 LRECL=54 GENERATED BY DATA

ALPHABETIC LIST OF VARIABLES

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT
3	ETA	NUM	8	30		
1	NOME	CHAR	25	4		
5	ORALE	NUM	8	46		
4	SCRITTO	NUM	8	38		
2	SESSO	CHAR	1	29		

----- SOURCE STATEMENTS -----

```
| DATA MEDIA.CLASSE3;  
| INPUT NOME $25.  
|     SESSO $ 27-27  
|     +2 ETA 2.  
|     @35 SCRITTO 35  
|     ,@40 ORALE 1.;  
| CARDS;
```

procedura contents.

```
PROC CONTENTS DATA=MEDIA ALL NOSOURCE;
```

SAS DATA SET DIRECTORY

NAME	# OBS
CLASSE2	21
CLASSE1	21
CLASSE3	21

CONTENTS OF SAS DATA SET MEDIA.CLASSE1

OBSERVATIONS=21 CREATED BY CMS USERID MESURE ON CPUID PP-3081-001964
AT 14:58 TUESDAY, APRIL 30, 1985 BY SAS RELEASE 82.3

FILE=MEDIA CLASSE2 BLKSIZE=8158 LRECL=54 GENERATED BY DATA

ALPHABETIC LIST OF VARIABLES

#	VARIABLE LABEL	TYPE	LENGTH	POSITION	FORMAT	INFORMAT
3	ETA	NUM	8	30		
1	NOME	CHAR	25	4		
5	ORALE	NUM	8	46		
4	SCRITTO	NUM	8	38		
2	SESSO	CHAR	1	29		

CONTENTS OF SAS DATA SET MEDIA.CLASSE2

OBSERVATIONS=21 CREATED BY CMS USERID MESURE ON CPUID PP-3081-001964
AT 14:58 TUESDAY, APRIL 30, 1985 BY SAS RELEASE 82.3

FILE=MEDIA CLASSE1 BLKSIZE=8158 LRECL=54 GENERATED BY DATA

ALPHABETIC LIST OF VARIABLES

#	VARIABLE LABEL	TYPE	LENGTH	POSITION	FORMAT	INFORMAT
3	ETA	NUM	8	30		
1	NOME	CHAR	25	4		
5	ORALE	NUM	8	46		
4	SCRITTO	NUM	8	38		
2	SESSO	CHAR	1	29		

procedura contents

```
PROC CONTENTS DATA=MEDIA.CLASSE3 SHORT;
```

CONTENTS OF SAS DATA SET MEDIA.CLASSE3

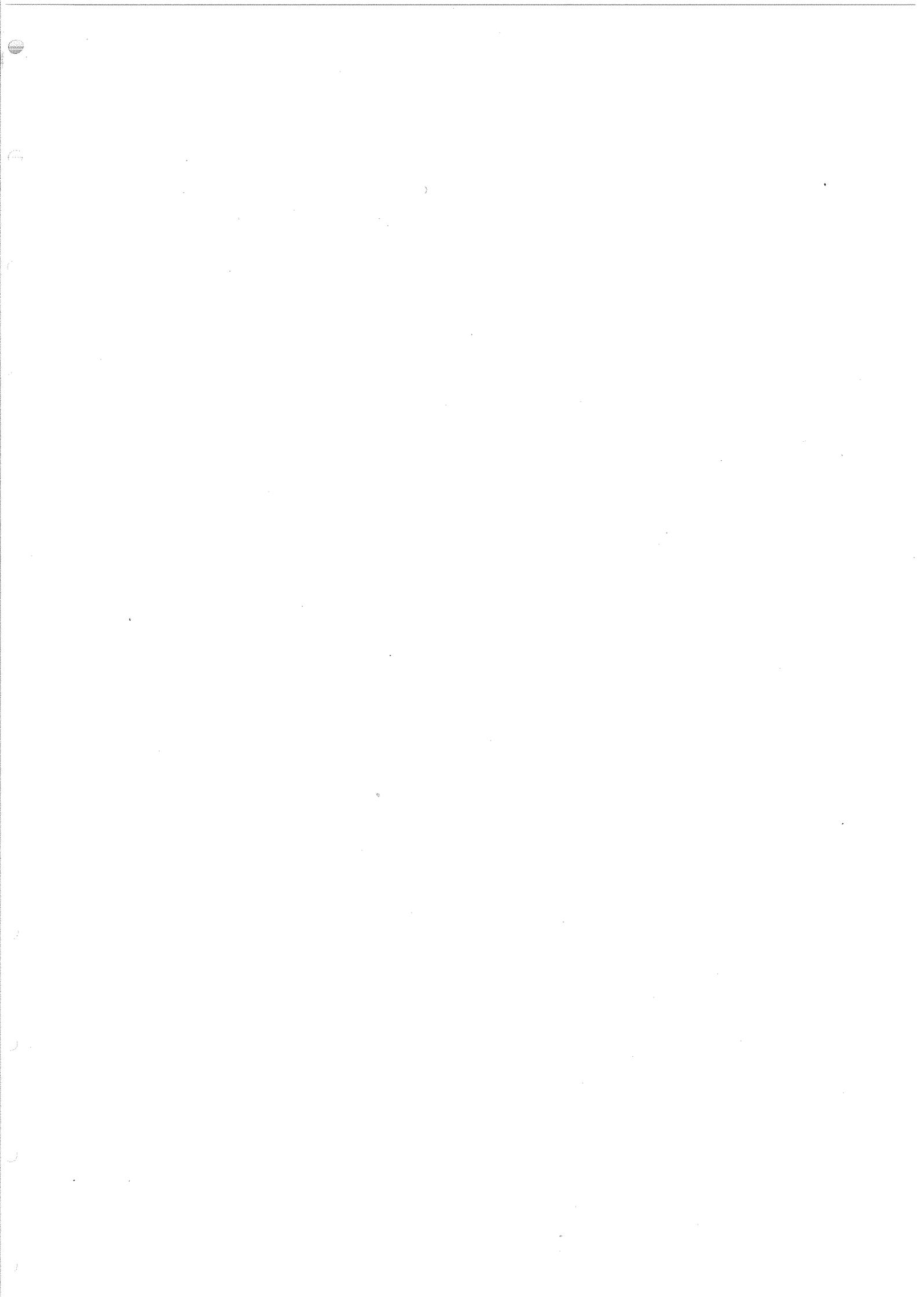
ALPHABETIC LIST OF VARIABLES

ETA NOME ORALE SCRITTO SESSO

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3;  
  BY NOME;  
PROC PRINT DATA=MEDIA.CLASSE3;
```

OBS	NOME	SESSO	ETA	SCRITTO	ORALE
1	ACERBI NILO	M	14	6	5
2	ANIA GIUSEPPE	M	15	4	4
3	BALLONI FRANCA	F	14	7	8
4	BELLANDI ANGELA	F	14	6	6
5	BERTINI MARIA	F	14	6	6
6	BIGOTTI MIRIANA	F	14	5	6
7	CORONA GIOVANNI	M	14	6	5
8	FANTONI CINZIA	F	13	6	6
9	GARZELLI PAOLO	M	14	.	.
10	GARZELLI VITTORIO	M	14	6	6
11	GHELARDI ANNA	F	14	6	7
12	GINI ALDO	M	15	6	6
13	LIPERINI FABRIZIO	M	15	.	5
14	MARRAZZO UMBERTO	M	14	5	6
15	MATTEI DINO	M	14	5	5
16	MICHELETTI BRUNO	M	14	6	5
17	HORI GINO	M	14	5	5
18	SORRI DANILO	M	13	7	6
19	TONINI RENZO	M	16	4	5
20	TOZZINI LOREDANA	F	14	5	6
21	VALLINI CLAUDIA	F	14	5	5



NELLE ESPRESSIONI LOGICHE (CONFRONTI) POSSONO ESSERE USATI

SEI OPERATORI :

FORMA 1	FORMA 2	SIGNIFICATO
<	LT	MINORE DI (LESS THAN)
>	GT	MAGGIORE DI (GREATER THAN)
=	EQ	UGUALE A (EQUAL THAN)
<=	LE	MINORE O UGUALE A (LESS THAN OR EQUAL TO)
>=	GE	MAGGIORE O UGUALE A (GREATER THAN OR EQUAL TO)
≠	NE	DIVERSO DA (NOT EQUAL TO)

GLI OPERATORI POSSONO ESSERE SCRITTI INDIFFERENTEMENTE NELLA FORMA 1 o 2 (NON TUTTI I TERMINALI O LE STAMPATRICI METTONO A DISPOSIZIONE I CARATTERI NECESSARI ALLA FORMA 1)

ESEMPIO:

IF ETA < 18 THEN RAGAZZI+1;

IF 18 <= ETA LT 75 THEN ADULTI+1;

IF ETA GE 75 THEN MATURI+1;

Istruzione DELETE

In SAS e' possibile controllare quali osservazione vogliamo che siano trascritte nel data set sas di uscita con l'istruzione DELETE.

Sintassi:

```
DELETE;
```

Esempio:

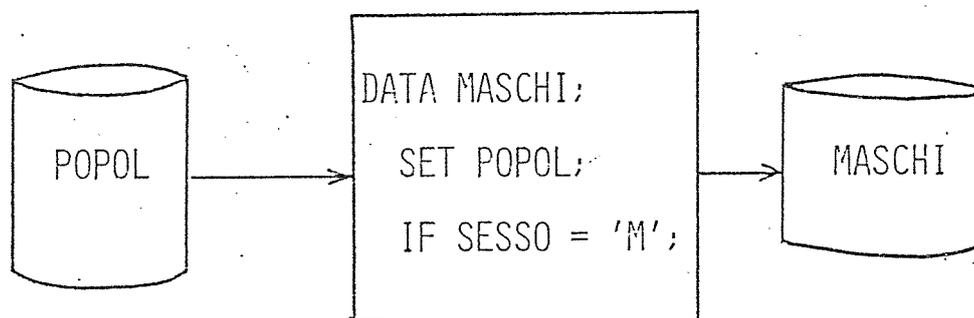
```
IF SESSO='M'  
  THEN DELETE;  
ELSE FEMMINE+1;
```

L'esecuzione di una DELETE provoca il trasferimento della esecuzione all'inizio dello step dopo una re-inizializzazione del PDU (House-keeping).

DICHIARAZIONE IF-SUBSETTING

UN MODO SEMPLICE DI INTRODURRE AL FUNZIONAMENTO DI UNA DICHIARAZIONE IF ^{SUBSETTING} È IL SEGUENTE:

SUPPONIAMO DI AVERE A DISPOSIZIONE UN ARCHIVIO DELLA POPOLAZIONE E DI VOLER SELEZIONARE DA ESSO TUTTI GLI INDIVIDUI MASCHI.



LA DICHIARAZIONE IF SESSO='M''; È CHIAMATA 'SUBSETTING IF';

IL MODO DI LAVORARE DI QUESTA DICHIARAZIONE RISULTA SIMILE AD UNA PORTA LA QUALE PERMETTE AD UN'OSSERVAZIONE DI "PASSARE" SOLO SE LA CONDIZIONE DELL'IF (SESSO='M') RISULTA VERA.

NEL CASO CHE ESSA RISULTI FALSA (IL SOGGETTO È UNA FEMMINA) ALLORA L'OSSERVAZIONE NON VIENE INSERITA NEL NUOVO ARCHIVIO MASCHI.

ESEMPI

```
DATA MASCHI;  
  SET POPOL;  
  IF SESSO <'M' THEN DELETE;
```

Si ottiene lo stesso risultato scrivendo:

```
DATA MASCHI;  
  SET POPOL;  
  IF SESSO='M';
```

IF-subsetting con condizione composta:

```
DATA DONNE_SP;  
  SET POPOL;  
  IF SESSO='F' AND STATO_CU='C';
```

Dichiarazioni DROP e KEEP

Tutte le variabili incontrate nello step di DATA sono presenti nel PDU.

Per controllare le variabili che saranno inserite nel data set SAS si usano la DROP e la KEEP.

In assenza di DROP/KEEP tutte le variabili del PDU sono aggiunte al data set SAS.

Sintassi:

DROP lista di variabili;

KEEP lista di variabili;

La DROP indica le variabili che NON DEVONO essere presenti nel data set SAS.

La KEEP indica le variabili che DEVONO essere presenti nel data set SAS.

ESEMPI

```
DATA ESEMPIO;  
  DROP A B C S;  
  D=A*B;  
  E=C;  
  CARDS;  
10 10 COSTANT1  
10 20 COSTANT2  
10 30 COSTANT3  
.....  
.....  
PROC PRINT;  
RUN;
```

PROGRAM DATA VECTOR



DATA SET A		
ENTRATE	USCITE	SALDO
2000000	1000000	1000000
3000000	500000	2500000

ESEMPIO:

DATA B;

SET A;

DROP ENTRATE USCITE;

VEETTORE DEI DATI

ENTRATE	USCITE	SALDO
---------	--------	-------

DATA SET B
SALDO
1000000
2500000

DA RIPORTARE NEL DATASET B

Dichiarazione DO iterativo

E' possibile far ripetere al SAS l'esecuzione di un gruppo di dichiarazioni tramite la DO.

Sintassi:

```
DO <indice> = <start> TO <stop> BY <step>;
```

```
DO <indice> = <start> TO <stop>  
  BY <step> WHILE <condiz.>,...;  
  UNTIL <condiz.>,...;
```

ESEMPIO:

```
DO I = 1 TO 100;  
  QUAD=I**2;  
  CUBO=I**3;  
  QUAD=I**4;  
END;
```

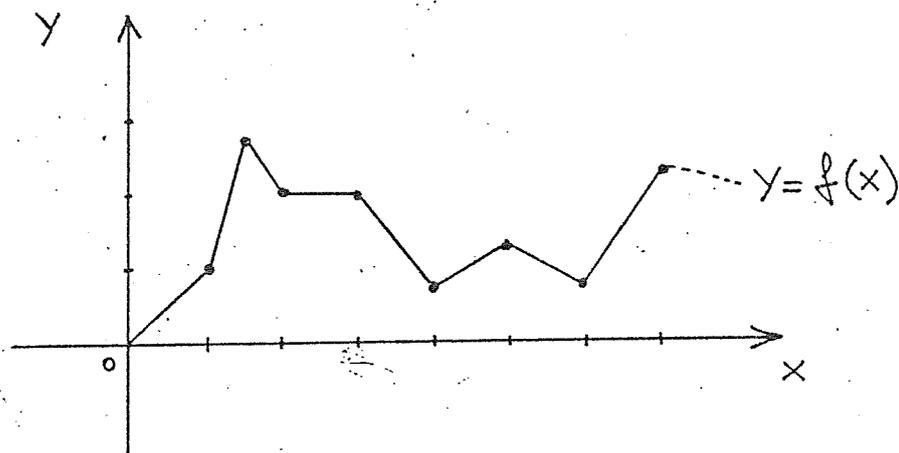
```
DO WHILE ( I>100 );  
  <Dichiaraz1>;  
  <Dichiaraz2>;  
  .....  
  <DichiarazN>;  
END;
```

```
DO I = 1,7,35;  
  <dichiaraz1>;  
  .....  
  <dichiarazN>;  
END;
```

```
DO UNTIL ( I>100 );  
  <Dichiaraz1>;  
  <Dichiaraz2>;  
  .....  
  <DichiarazN>;  
END;
```

PROCEDURA PLOT

LA NECESSITÀ DI INDIVIDUARE L'EVOLVERSI (L'ANDAMENTO, IL TREND) DI UN FENOMENO (ES. NASCITE, CRISI, ELEZIONI, ECC.) CI PORTA IN MOLTE OCCASIONI A DISEGNARE DEI GRAFICI DEL TIPO:



AL FINE DI OTTENERE DISEGNI DEL GENERE, IL SAS METTE A DISPOSIZIONE LA PROCEDURA PLOT.

SINTASSI:

```
PROC PLOT OPZIONI DATA=NOME ARCHIVIO SAS;
```

ESEMPIO D'USO DELLA PROC PLOT

DESIDERIAMO TRACCIARE IL GRAFICO DELLA SEGUENTE EQUAZIONE (RETTA):

$$Y = 2 X + 25$$

FACENDO VARIARE LA VARIABILE INDIPENDENTE X DA 0 A 100 SECONDO INCREMENTI COSTANTI DI 5.

IL PROGRAMMA È IL SEGUENTE:

```

A ← { DATA PUNTI;
      DO X=0 TO 100 BY 5;
      Y=2*X+25;
      OUTPUT;
      END;

B ← { PROC PLOT;
      PLOT Y*X='x';
      TITLE GRAFICO DI Y=2X+25;
    }

```

GRAFICO DI $Y=2X+25$

PLOT OF $Y=2X$ SYMBOL USED IS *

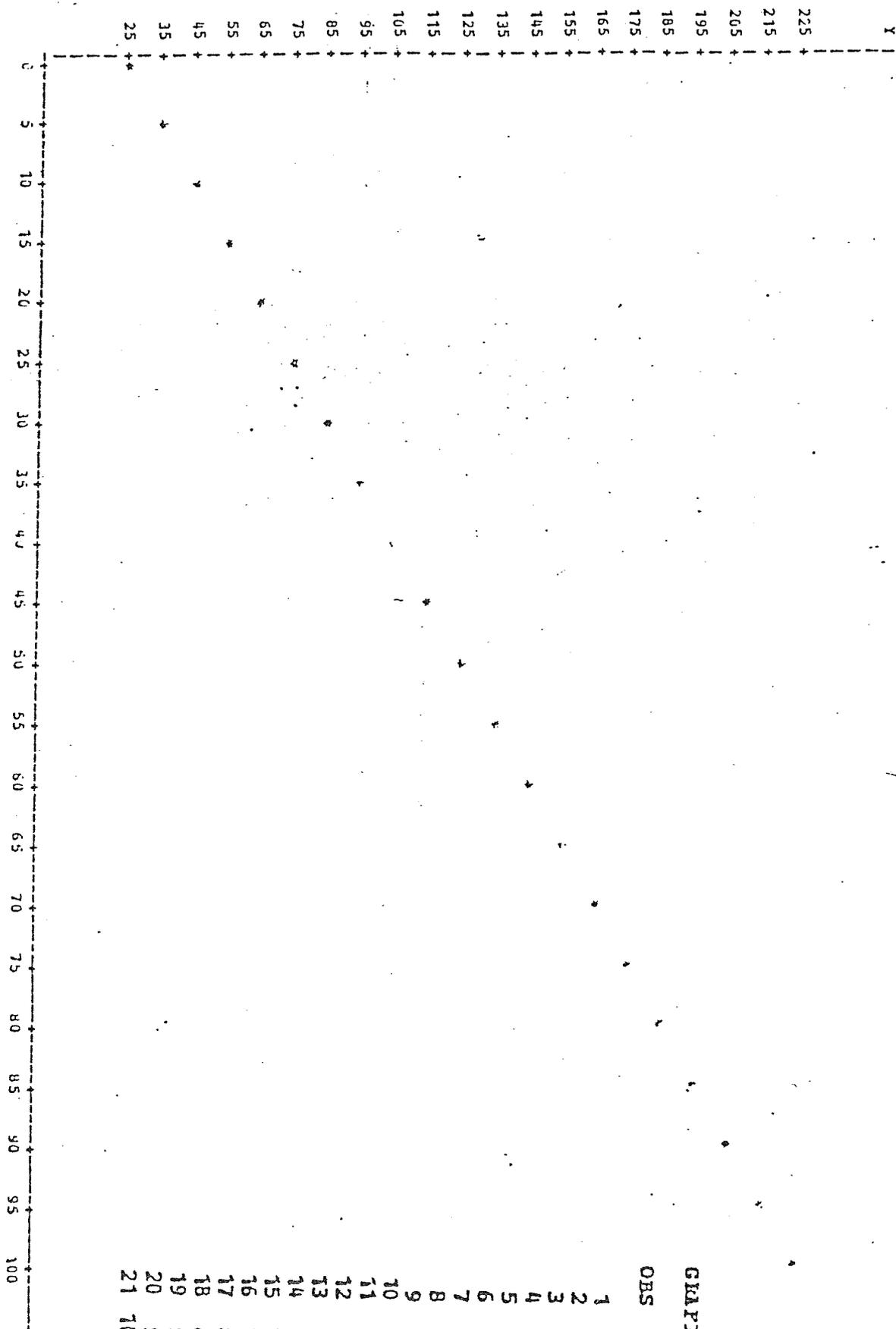


GRAFICO DI $Y=2X+25$

OBS	X	Y
1	0	25
2	5	35
3	10	45
4	15	55
5	20	65
6	25	75
7	30	85
8	35	95
9	40	105
10	45	115
11	50	125
12	55	135
13	60	145
14	65	155
15	70	165
16	75	175
17	80	185
18	85	195
19	90	205
20	95	215
21	100	225

Alcuni cenni alle funzioni SAS

Il SAS mette a disposizione una serie di funzioni (functions) che possono essere utilizzate nelle espressioni numeriche e non, come operatori speciali. Una funzione opera su argomenti e fornisce in uscita un valore.

Le funzioni che operano su argomenti esclusivamente numerici si dicono di tipo numerico, viceversa di tipo carattere. Esistono anche funzioni che elaborano date, tempi o date e tempi; esistono funzioni trigonometriche, probabilistiche, statistiche, ecc.

Sintassi:

```
NOME-FUNZIONE( argom1, argom2, ..., argomN );  
NOME-FUNZIONE( OF variab1-variabN );  
NOME-FUNZIONE( OF variab variab variab );
```

Esempi di funzioni numeriche:

```
Y=ABS(X);  
TOTSTIP=SUM(OF STIPEN1-STIPEN12);  
MINORE=MIN(X Y Z);  
RADQUAD=SQRT(NUMERO);  
MEDIA=MEAN(OF X1-X160);
```

Esempi di funzioni carattere:

```
CODICE='A360';  
PNUMER=SUBSTR(CODICE,1,1);
```

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
  BY SESSO;  
PROC PRINT DATA=SORT.CLASSE3;
```

OBS	NOME	SESSO	ETA	SCRITTO	ORALE
1	TOZZINI LOREDANA	F	14	5	6
2	BERTINI MARIA	F	14	6	6
3	GHELARDI ANNA	F	14	6	7
4	BIGOTTI MIRIANA	F	14	5	6
5	FANTONI CINZIA	F	13	6	6
6	BALLONI FRANCA	F	14	7	8
7	VALLINI CLAUDIA	F	14	5	5
8	BELLANDI ANGELA	F	14	6	6
9	MORI GINO	M	14	5	5
10	SORRI DANILO	M	13	7	6
11	LIPERINI FABRIZIO	M	15	-	5
12	GAEZELLI PAOLO	M	14	-	-
13	ANIA GIUSEPPE	M	15	4	4
14	MICHELETTI BRUNO	M	14	6	5
15	GARZELLI VITTORIO	M	14	6	6
16	TONINI RENZO	M	16	4	5
17	ACERBI NILO	M	14	6	5
18	GINI ALDO	M	15	6	6
19	CORONA GIOVANNI	M	14	6	5
20	MATTEI DINO	M	14	5	5
21	MARRAZZO UMBERTO	M	14	5	6

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
  BY SESSO NOME;  
PROC PRINT DATA=SORT.CLASSE3;
```

OBS	NOME	SESSO	ETA	SCRITTO	ORALE
1	BALLONI FRANCA	F	14	7	8
2	BELLANDI ANGELA	F	14	6	6
3	BERTINI MARIA	F	14	6	6
4	BIGOTTI MIRIANA	F	14	5	6
5	FANTONI CINZIA	F	13	6	6
6	GHELARDI ANNA	F	14	6	7
7	TOZZINI LOREDANA	F	14	5	6
8	VALLINI CLAUDIA	F	14	5	5
9	ACERBI NILO	M	14	6	5
10	ANIA GIUSEPPE	M	15	4	4
11	CORONA GIOVANNI	M	14	6	5
12	GARZELLI PAOLO	M	14	.	.
13	GARZELLI VITTORIO	M	14	6	6
14	GINI ALDO	M	15	6	6
15	LIPERINI FABRIZIO	M	15	.	5
16	MARRAZZO UMBERTO	M	14	5	6
17	MATTEI DINO	M	14	5	5
18	MICHELETTI BRUNO	M	14	6	5
19	MORI GINO	M	14	5	5
20	SOERI DANILO	M	13	7	6
21	TONINI RENZO	M	16	4	5

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
  BY SESSO DESCENDING ETA HOME;  
PROC PRINT DATA=SORT.CLASSE3;
```

OBS	NOME	SESSO	ETA	SCRITTO	ORALE
1	BALLONI FRANCA	F	14	7	8
2	BELLANDI ANGELA	F	14	6	6
3	BERTINI MARIA	F	14	6	6
4	BIGOTTI MIRIANA	F	14	5	6
5	GHELARDI ANNA	F	14	6	7
6	TOZZINI LOREDANA	F	14	5	6
7	VALLINI CLAUDIA	F	14	5	5
8	FANTONI CINZIA	F	13	6	6
9	TONINI RENZO	M	16	4	5
10	ANIA GIUSEPPE	M	15	4	4
11	GINI ALDO	M	15	6	6
12	LIPERINI FABRIZIO	M	15	.	5
13	ACERBI NILO	M	14	6	5
14	CORONA GIOVANNI	M	14	6	5
15	GARZELLI PAOLO	M	14	.	.
16	GARZELLI VITTORIO	M	14	6	6
17	MARRAZZO UMBERTO	M	14	5	6
18	MATTEI DIINO	M	14	5	5
19	MICHELETTI BRUNO	M	14	6	5
20	MORI GINO	M	14	5	5
21	SORRI DANILO	M	13	7	6

procedura print

```
PROC PRINT DATA=MEDIA.CLASSE3 N;  
TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';  
VAR NOME ETA;
```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

OBS	NOME	ETA
1	MORI GINO	14
2	SORRI DANILO	13
3	LIPERINI FABRIZIO	15
4	TOZZINI LOREDANA	14
5	BERTINI MARIA	14
6	GARZELLI PAOLO	14
7	GHELARDI ANNA	14
8	ANIA GIUSEPPE	15
9	KICHELETTI BRUNO	14
10	GARZELLI VITTORIO	14
11	BIGOTTI MIRIANA	14
12	TONINI RENZO	16
13	FANTONE CINZIA	13
14	BALLONI FRANCA	14
15	ACERBI NILO	14
16	GINI ALDO	15
17	VALLINI CLAUDIA	14
18	CORONA GIOVANNI	14
19	MATTEI DINO	14
20	BELLANDI ANGELA	14
21	MARRAZZO UMBERTO	14

N=21

procedura print

```
PROC PRINT DATA=MEDIA,CLASSE3,DOUBLE;  
TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';  
VAR ETA SCRITTO;  
ID NOME;
```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

NOME	ETA	SCRITTO
MORI GINO	14	5
SORRI DANILO	13	7
LIPERINI FABRIZIO	15	.
TOZZINI LOREDANA	14	5
BERTINI MARIA	14	6
GARZELLI PAOLO	14	.
GHELARDI ANNA	14	6
ANIA GIUSEPPE	15	4
MICHELETTI BRUNO	14	6
GARZELLI VITTORIO	14	6
BIGOTTI MIRIANA	14	5
TONINI RENZO	16	4
FANTONI CINZIA	13	6
BALLONI FRANCA	14	7
ACERBI NILO	14	6
GINI ALDO	15	6
VALLINI CLAUDIA	14	5
CORONA GIOVANNI	14	6
HATTEI DINO	14	5
BELLANDI ANGELA	14	6
MARRAZZO UMBERTO	14	5

procedura print

```

PROC SORT DATA=KEDIA.CLASSE3 OUT=SORT.CLASSE3;
  BY SESSO;
PROC PRINT DATA=SORT.CLASSE3 UNIFORM SPLIT=7;
  TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';
  LABEL NOBE = 'COGNOME E NOME';
       SCRITTO = 'VOTO SCRITTO/DI ITALIANO';
  ID NOBE;
  BY SESSO;
  
```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

SESSO=F

COGNOME E NOME	ETA	VOTO SCRITTO DI ITALIANO	ORALE
BALLONI FRANCA	14	7	8
BELLANDI ANGELA	14	6	6
BERTINI MARIA	14	6	6
BIGOTTI MIRIANA	14	5	6
GHELARDI ANNA	14	6	7
TOZZINI LOREDANA	14	5	6
VALLINI CLAUDIA	14	5	5
FANTONI CINZIA	13	6	6

SESSO=M

COGNOME E NOME	ETA	VOTO SCRITTO DI ITALIANO	ORALE
TONINI RENZO	16	4	5
ANIA GIUSEPPE	15	4	4
GINI ALDO	15	6	6
LIPERINI FABRIZIO	15	.	5
ACERBI NILO	14	6	5
CORONA GIOVANNI	14	6	5
GARZELLI PAOLO	14	.	.
GARZELLI VITTORIO	14	6	6
MARRAZZO UMBERTO	14	5	6
MATTEI DINO	14	5	5
MICHELETTI BRUNO	14	6	5
MORI GINO	14	5	5
SORRI DANILO	13	7	6

procedura print

```

PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;
  BY SESSO;
PROC PRINT DATA=SORT.CLASSE3 SPLIT=/:
  TITLE 'SOMMA DEI VOTI DI ITALIANO DELLA CLASSE III A';
  LABEL NONE = 'COGNOME E NOME'
        ETA = 'ETA'
        SCRITTO = 'VOTO SCRITTO/DI ITALIANO'
        ORALE = 'VOTO ORALE/DI ITALIANO';
  ID NONE;
  BY SESSO;
  SUM SCRITTO ORALE;

```

SOMMA DEI VOTI DI ITALIANO DELLA CLASSE III A

SESSO=F

COGNOME E NOME	ETA'	VOTO SCRITTO DI ITALIANO	VOTO ORALE DI ITALIANO
BALLONI FRANCA	14	7	8
BELLANDI ANGELA	14	6	6
BERTINI MARIA	14	6	6
BIGOTTI MIRIANA	14	5	6
GHELARDI ANNA	14	6	7
TOZZINI LOREDANA	14	5	6
VALLINI CLAUDIA	14	5	5
FANTONI CINZIA	13	6	6
-----		46	50
SESSO			

SESSO=M

COGNOME E NOME	ETA'	VOTO SCRITTO DI ITALIANO	VOTO ORALE DI ITALIANO
TONINI RENZO	16	4	5
ANIA GIUSEPPE	15	4	4
GINI ALDO	15	6	6
LIPERINI FABRIZIO	15	-	5
ACERBI NILO	14	6	5
CORONA GIOVANNI	14	6	5
GARZELLI PAOLO	14	-	-
GARZELLI VITTORIO	14	6	6
MARRAZZO UMBERTO	14	5	6
MATTEI DINO	14	5	5
MICHELETTI BRUNO	14	6	5
MORI GINO	14	5	5
SORRI DANILO	13	7	6
-----		60	63
SESSO		===	===
		106	113

procedura format

```

PROC FORMAT;
  VALUE $SEX
    F = 'FEMMINA';
    M = 'MASCHIO';
    OTHER = 'NEUTRO';
  VALUE VOTO
    . = 'NON CLASSIFICATO';
    0-5 = 'INSUFFICIENTE';
    6,7 = 'SUFFICIENTE';
    8-HIGH = 'BUONO';
  PICTURE ANNI
    LOW-HIGH = '0999';
PROC PRINT DATA=MEDIA.CLASSE3 SPLIT=7;
  TITLE 'ELENCO DEGLI ALUNNI DELLA CLASSE III A';
  LABEL COGNOME = 'COGNOME E NOME';
  ETA = 'ETA';
  SCRITTO = 'VOTO SCRITTO/DI ITALIANO';
  ORALE = 'VOTO ORALE/DI ITALIANO';
  FORMAT $SEX.;
  FORMAT SCRITTO VOTO.;
  FORMAT ORALE VOTO.;
  FORMAT ANNI.;
  ID NOME;

```

ELENCO DEGLI ALUNNI DELLA CLASSE III A

COGNOME E NOME	SESSO	ETA'	VOTO SCRITTO DI ITALIANO	VOTO ORALE DI ITALIANO
BALLONI FRANCA	FEMMINA	014	SUFFICIENTE	BUONO
BELLANDI ANGELA	FEMMINA	014	SUFFICIENTE	SUFFICIENTE
BERTINI MARIA	FEMMINA	014	SUFFICIENTE	SUFFICIENTE
BIGOTTI MIRIANA	FEMMINA	014	INSUFFICIENTE	SUFFICIENTE
GHELARDI ANNA	FEMMINA	014	SUFFICIENTE	SUFFICIENTE
TOZZINI LOREDANA	FEMMINA	014	INSUFFICIENTE	SUFFICIENTE
VALLINI CLAUDIA	FEMMINA	014	INSUFFICIENTE	INSUFFICIENTE
FANTONI CINZIA	FEMMINA	013	SUFFICIENTE	SUFFICIENTE
TONINI RENZO	MASCHIO	016	INSUFFICIENTE	INSUFFICIENTE
ANIA GIUSEPPE	MASCHIO	015	INSUFFICIENTE	INSUFFICIENTE
GINI ALDO	MASCHIO	015	SUFFICIENTE	SUFFICIENTE
LIPERINI FABRIZIO	MASCHIO	015	NON CLASSIFICATO	INSUFFICIENTE
ACERBI NILO	MASCHIO	014	SUFFICIENTE	INSUFFICIENTE
CORONA GIOVANNI	MASCHIO	014	SUFFICIENTE	INSUFFICIENTE
GARZELLI PAOLO	MASCHIO	014	NON CLASSIFICATO	NON CLASSIFICATO
GARZELLI VITTORIO	MASCHIO	014	SUFFICIENTE	SUFFICIENTE
MARRAZZO UMBERTO	MASCHIO	014	INSUFFICIENTE	SUFFICIENTE
MATTEI DINO	MASCHIO	014	INSUFFICIENTE	INSUFFICIENTE
MICHELETTI BRUNO	MASCHIO	014	SUFFICIENTE	INSUFFICIENTE
MORI GINO	MASCHIO	014	INSUFFICIENTE	INSUFFICIENTE
SORRI DANILO	MASCHIO	013	SUFFICIENTE	SUFFICIENTE

procedura means

PROC MEANS DATA=MEDIA CLASSE3;

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD. ERROR OF MEAN
ETA	21	14.142857	0.654653	13.000000	16.000000	0.142857
SCRITTO	19	5.578947	0.837707	4.000000	7.000000	0.192183
ORALE	20	5.650000	0.875093	4.000000	8.000000	0.195676

procedura means

```
PROC MEANS DATA=MEDIA CLASSES MAXDEC=2;  
VAR SCRITTO ORALE;
```

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN
SCRITTO	19	5.58	0.84	4.00	7.00	0.19
ORALE	20	5.65	0.88	4.00	8.00	0.20

procedura means

```
PROC MEANS DATA=MEDIA.CLASSE3 MAXDEC=2 N SUM RANGE MEAN VAR NMISS;  
TITLE 'MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A';  
VAR SCRITTO ORALE;
```

MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A

VARIABLE	N	SUM	RANGE	MEAN	VARIANCE	N MISSING
SCRITTO	19	106.00	3.00	5.58	0.70	2
ORALE	20	113.00	4.00	5.65	0.77	1

procedura means

```

PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;
  BY SESSO;
PROC MEANS DATA=SORT.CLASSE3 MAXDEC=2;
  TITLE 'MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A';
  VAR SCRITTO ORALE;
  BY SESSO;
    
```

MEDIA DEI VOTI DI ITALIANO DELLA CLASSE III A

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN

SESSO=F						
SCRITTO	8	5.75	0.71	5.00	7.00	0.25
ORALE	8	6.25	0.89	5.00	8.00	0.31

SESSO=M						
SCRITTO	11	5.45	0.93	4.00	7.00	0.28
ORALE	12	5.25	0.62	4.00	6.00	0.18

procedura freq

PROC FREQ DATA=MEDIA.CLASSE3:

NOME	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
ACERBI NILO	1	1	4.762	4.762
ANIA GIUSEPPE	1	2	4.762	9.524
BALLONI FRANCA	1	3	4.762	14.286
BELLANDI ANGELA	1	4	4.762	19.048
BERTINI MARIA	1	5	4.762	23.810
BIGOTTI MIRIANA	1	6	4.762	28.571
CORONA GIOVANNI	1	7	4.762	33.333
FANTONI CINZIA	1	8	4.762	38.095
GARZELLI PAOLO	1	9	4.762	42.857
GARZELLI VITTORI	1	10	4.762	47.619
GHELARDI ANNA	1	11	4.762	52.381
GINI ALDO	1	12	4.762	57.143
LIPERINI FABRIZI	1	13	4.762	61.905
MARRAZZO UMBERTO	1	14	4.762	66.667
MATTEI DINO	1	15	4.762	71.429
MICHELETTI BRUNO	1	16	4.762	76.190
MORI GINO	1	17	4.762	80.952
SORRI DANILO	1	18	4.762	85.714
TONINI RENZO	1	19	4.762	90.476
TOZZINI LOREDANA	1	20	4.762	95.238
VALLINI CLAUDIA	1	21	4.762	100.000

SESSO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
F	8	8	38.095	38.095
M	13	21	61.905	100.000

ETA	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
13	2	2	9.524	9.524
14	15	17	71.429	80.952
15	3	20	14.286	95.238
16	1	21	4.762	100.000

SCRITTO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
-	2	-	-	-
4	2	2	10.526	10.526
5	6	8	31.579	42.105
6	9	17	47.368	89.474
7	2	19	10.526	100.000

ORALE	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
-	1	-	-	-
4	1	1	5.000	5.000
5	8	9	40.000	45.000
6	9	18	45.000	90.000
7	1	19	5.000	95.000
8	1	20	5.000	100.000

procedura freq

```
PROC FREQ DATA=MEDIA.CLASSE3 ORDER=FREQ;  
TITLE "FREQUENZA DELLE VARIABILI";  
TABLE SESSO ETA;
```

FREQUENZA DELLE VARIABILI

SESSO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
M	13	13	61.905	61.905
F	8	21	38.095	100.000

ETA	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
14	15	15	71.429	71.429
15	3	18	14.286	85.714
13	2	20	9.524	95.238
16	1	21	4.762	100.000

procedura freq

```

PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;
  BY SESSO;
PROC FREQ DATA=SORT.CLASSE3 ORDER=FREQ;
  TITLE 'FREQUENZA DELLE VARIABILI';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO'
        ORALE = 'VOTO ORALE DI ITALIANO';
  TABLE SCRITTO;
  BY SESSO;
  
```

FREQUENZA DELLE VARIABILI SESSO=F

VOTO SCRITTO DI ITALIANO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
6	4	4	50.000	50.000
5	3	7	37.500	87.500
7	1	8	12.500	100.000

FREQUENZA DELLE VARIABILI SESSO=M

VOTO SCRITTO DI ITALIANO	FREQUENCY	CUM FREQ	PERCENT	CUM PERCENT
.	2	2	45.455	45.455
6	5	7	27.273	72.727
5	3	10	18.182	90.909
4	2	12	9.091	100.000
7	1	13		

procedura freq

```

PROC FREQ DATA=MEDIA.CLASSE3 ORDER=FREQ;
  TITLE 'FREQUENZA DELLE VARIABILI';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO'
        ORALE = 'VOTO ORALE DI ITALIANO';
  TABLE SESSO*SCRITTO;
  
```

FREQUENZA DELLE VARIABILI

TABLE OF SESSO BY SCRITTO

SESSO SCRITTO VOTO SCRITTO DI ITALIANO

FREQUENCY						TOTAL
PERCENT						
ROW PCT						
COL PCT	6	5	7	4		
M	2	5	3	1	2	11
	26.32	15.79	5.26	10.53		57.89
	45.45	27.27	9.09	18.18		
	55.56	50.00	50.00	100.00		
F	0	4	3	1	0	8
	21.05	15.79	5.26	0.00		42.11
	50.00	37.50	12.50	0.00		
	44.44	50.00	50.00	0.00		
TOTAL	9	6	2	2		19
	47.37	31.58	10.53	10.53		100.00

procedura freq

```

PROC FREQ DATA=MEDIA.CLASSE3 ORDER=FREQ;
  TITLE 'FREQUENZA DELLE VARIABILI';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO'
        ORALE = 'VOTO ORALE DI ITALIANO';
  TABLE SESSO*(SCRITTO ORALE)/NOFREQ NOROW NOCOL;
  
```

FREQUENZA DELLE VARIABILI

TABLE OF SESSO BY SCRITTO

SESSO	SCRITTO	VOTO SCRITTO DI ITALIANO				TOTAL
PERCENT		6	5	7	4	
M		26.32	15.79	5.26	10.53	57.89
F		21.05	15.79	5.26	0.00	42.11
TOTAL		9	6	2	2	19
		47.37	31.58	10.53	10.53	100.00

FREQUENZA DELLE VARIABILI

TABLE OF SESSO BY ORALE

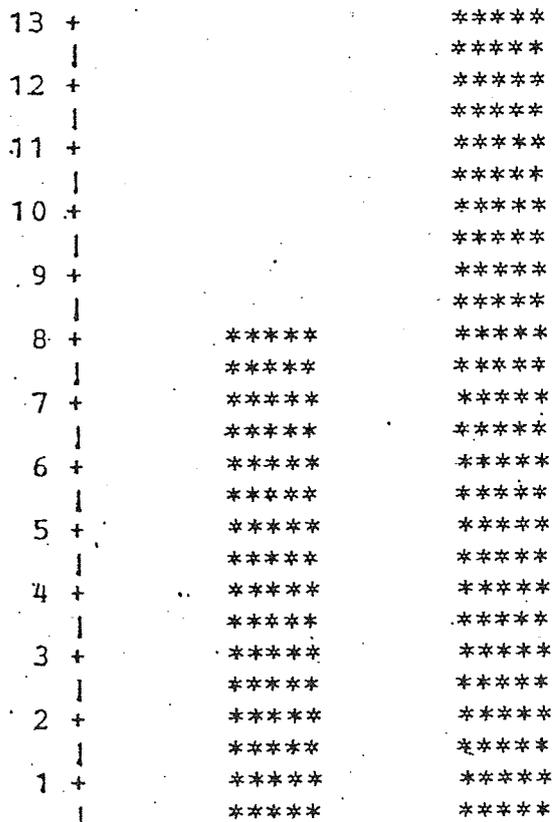
SESSO	ORALE	VOTO ORALE DI ITALIANO				TOTAL	
PERCENT		6	5	8	7	4	
M		20.00	35.00	0.00	0.00	5.00	60.00
F		25.00	5.00	5.00	5.00	0.00	40.00
TOTAL		9	8	1	1	1	20
		45.00	40.00	5.00	5.00	5.00	100.00

procedura chart

```
PROC CHART DATA=MEDIA.CLASSE3;  
VBAR SESSO / TYPE=SUM;
```

FREQUENCY BAR CHART

FREQUENCY



F M

SESSO

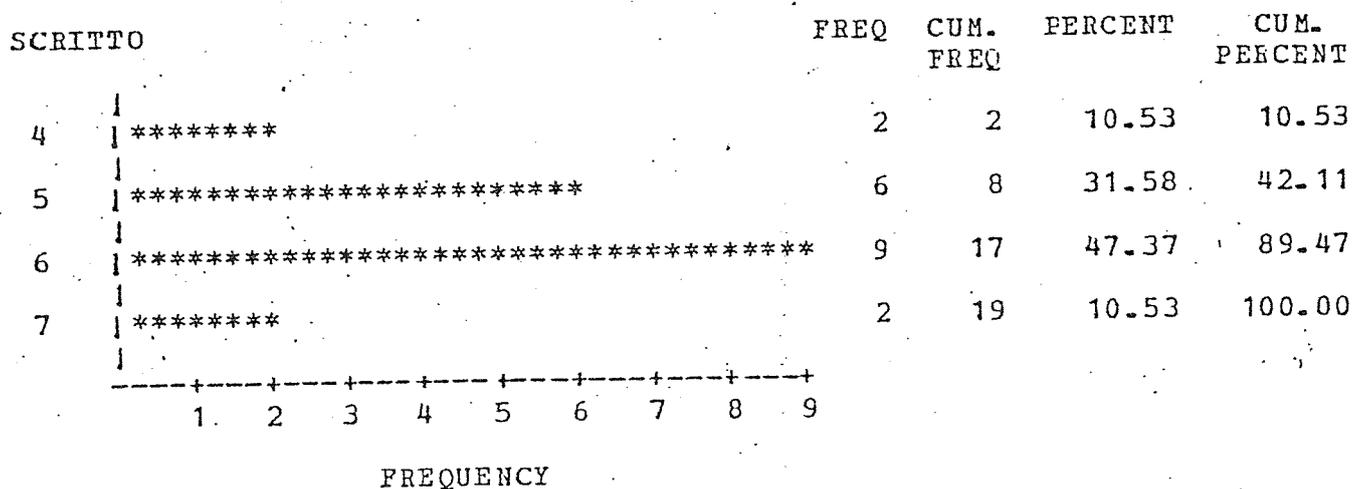
procedura chart

```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';
  ORALE = 'VOTO ORALE DI ITALIANO';
  EBAN SCRITTO ORALE / TYPE=FREQ DISCRETE;
  
```

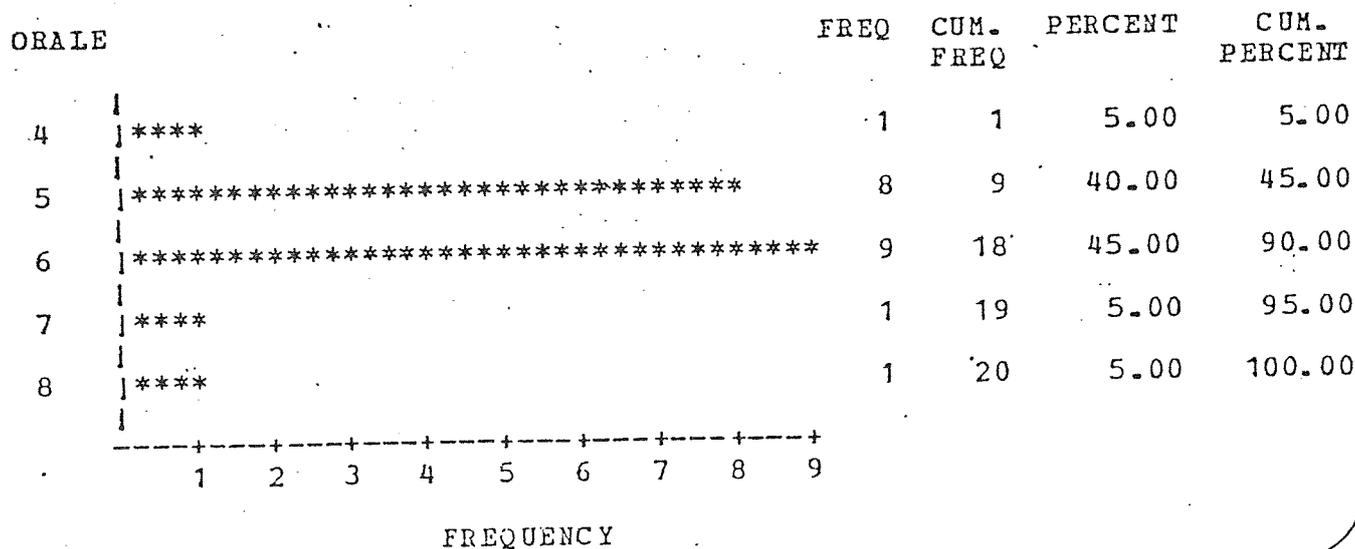
DISTRIBUZIONE DEI VOTI DI ITALIANO

FREQUENCY BAR CHART



DISTRIBUZIONE DEI VOTI DI ITALIANO

FREQUENCY BAR CHART



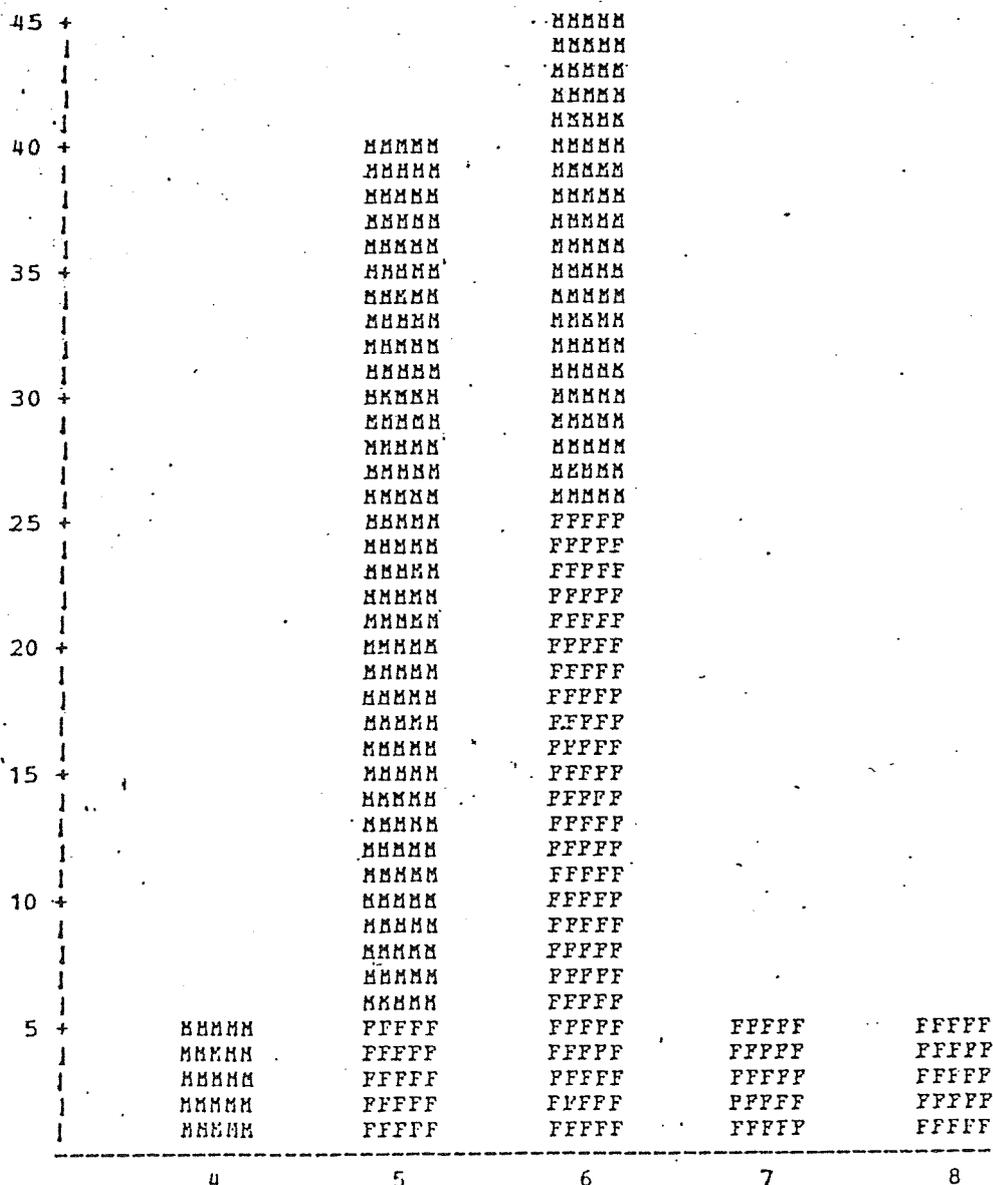
procedura chart

```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL ORALE = 'VOTO ORALE DI ITALIANO';
  VBAR ORALE / TYPE=PCT DISCRETE SUBGROUP=SESSO;
  
```

PERCENTAGE BAR CHART

PERCENTAGE



ORALE

procedura chart

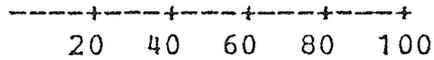
```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';
  HBAR SCRITTO / TYPE=CPCT DISCRETE GROUP=SESSO;
  
```

DISTRIBUZIONE DEI VOTI DI ITALIANO

CUMULATIVE PERCENTAGE BAR CHART

SESSO	SCRITTO	FREQ	CUM. FREQ	PERCENT	CUM. PERCENT
F	4	0	0	0.00	0.00
	5	3	3	15.79	15.79
	6	4	7	21.05	36.84
	7	1	8	5.26	42.11
M	4	2	10	10.53	52.63
	5	3	13	15.79	68.42
	6	5	18	26.32	94.74
	7	1	19	5.26	100.00



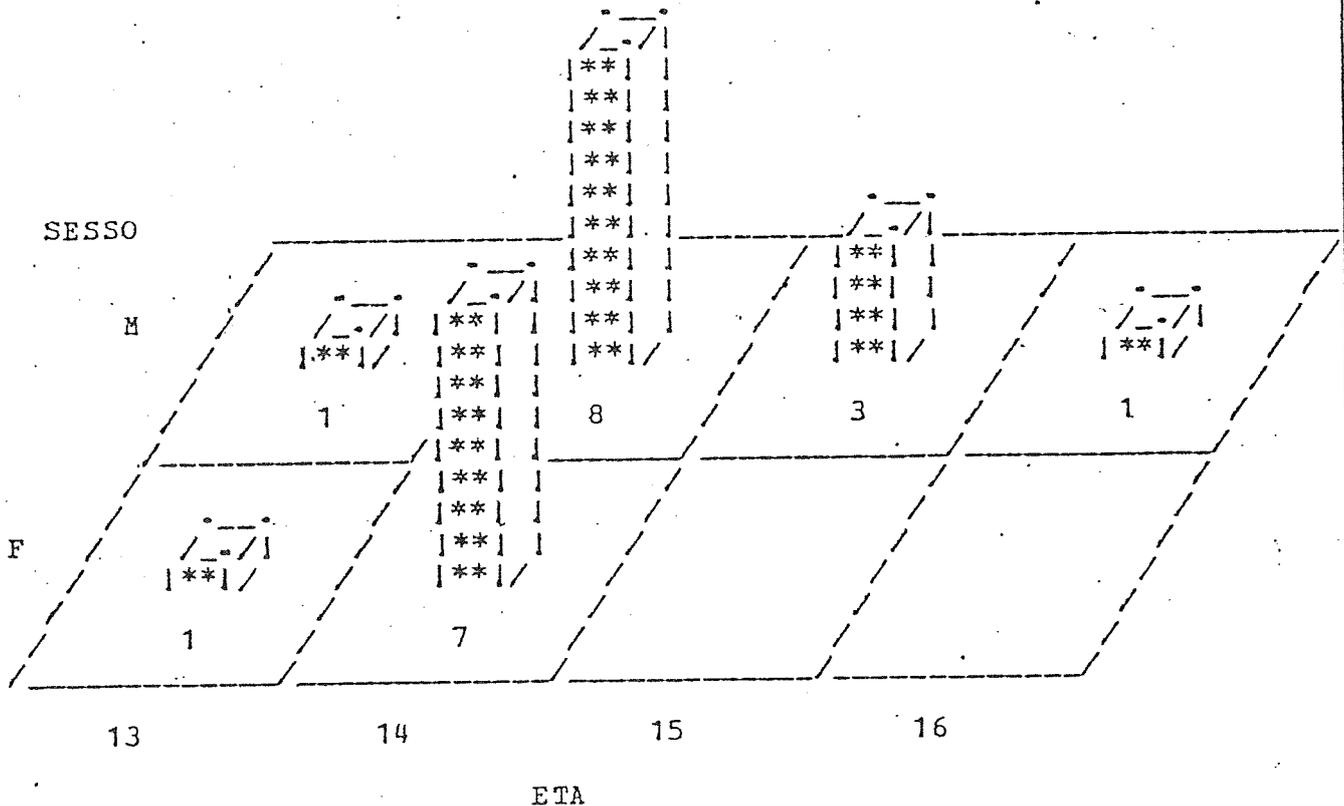
CUMULATIVE PERCENTAGE

procedura chart

```

PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DELLE ETA';
  LABEL ETA = 'ETA';
  BLOCK ETA / TYPE=FRFQ DISCRETE GROUP=SESSO;
  
```

DISTRIBUZIONE DELLE ETA
FREQUENCY BLOCK CHART



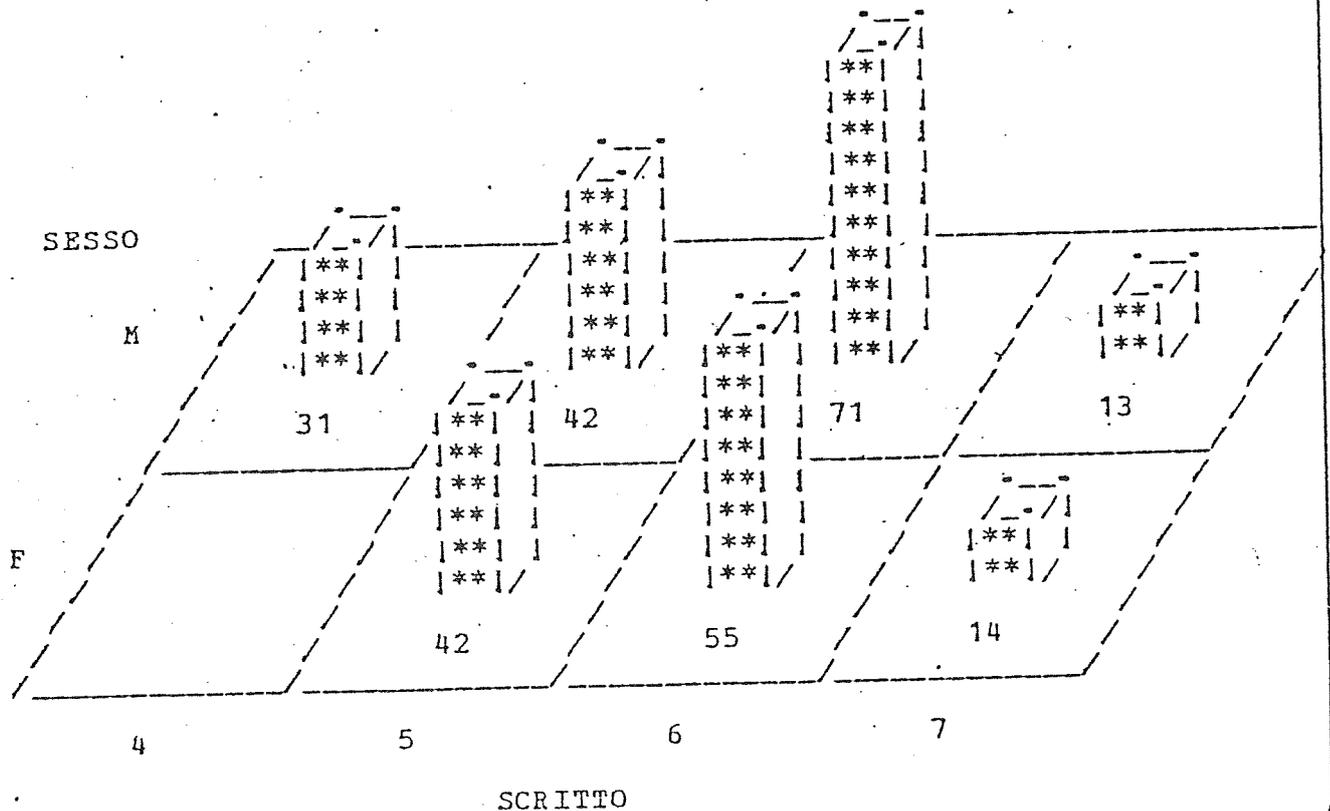
procedura chart

```

PROC CHART DATA=MEDIA-CLASSE;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';
  LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';
  BLOCK SCRITTO / TYPE=REQ DISCRETE GROUP=SESSO SUMVAR=ETA;
  
```

DISTRIBUZIONE DEI VOTI DI ITALIANO

BLOCK CHART OF ETA SUMS

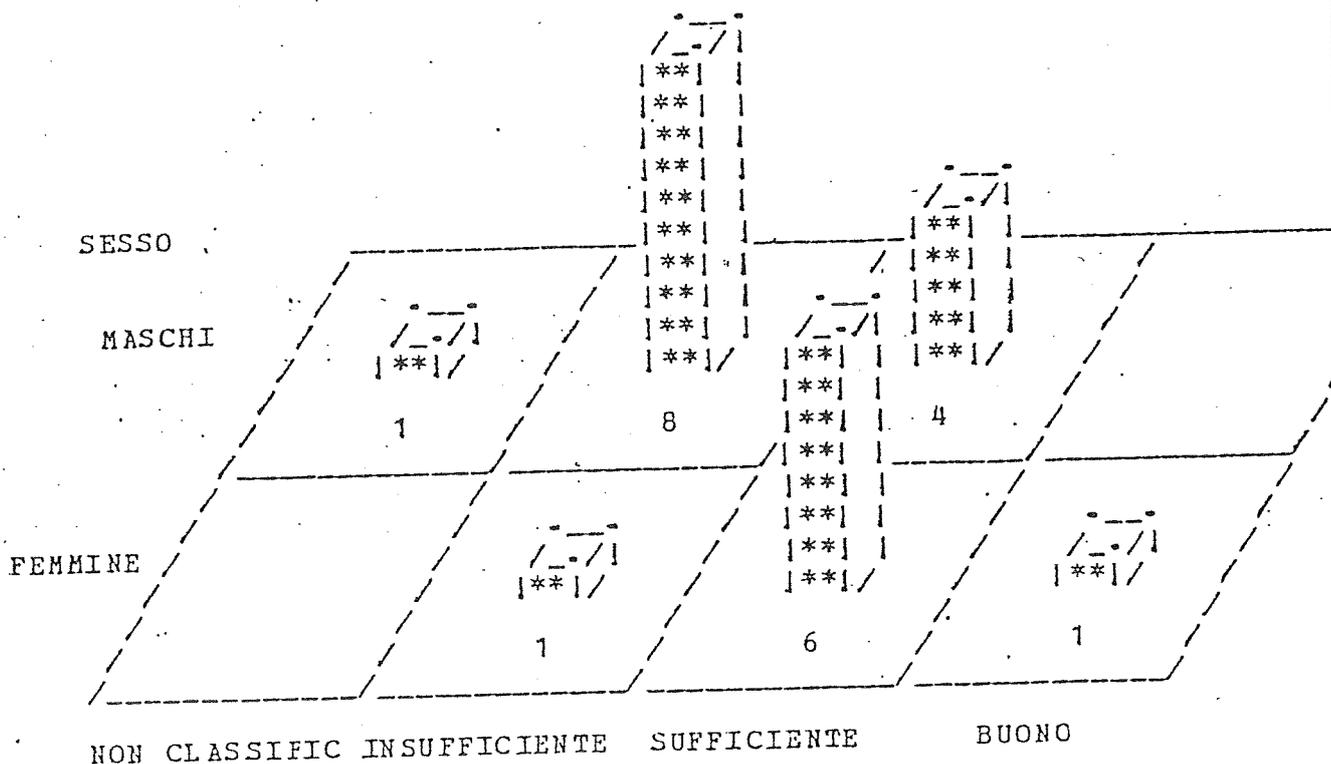


procedura format

```

PROC FORMAT;
  VALUE $SEX
    F = 'FEMMINE'
    M = 'MASCHI'
    OTHER = 'NEUTRO';
  VALUE VOTO
    - = 'NON CLASSIFICATO'
    0-5 = 'INSUFFICIENTE'
    6,7 = 'SUFFICIENTE'
    8-HIGH = 'BUONO';
PROC CHART DATA=MEDIA.CLASSE3;
  TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO DELLA CLASSE III A';
  LABEL ORALE = 'VOTO ORALE DI ITALIANO';
  FORMAT SESSO $SEX;
  FORMAT ORALE VOTO;
  BLOCK ORALE / DISCRETE TYPE=FREQ GROUP=SESSO MISSING;
  
```

DISTRIBUZIONE DEI VOTI DI ITALIANO DELLA CLASSE III A
 FREQUENCY BLOCK CHART

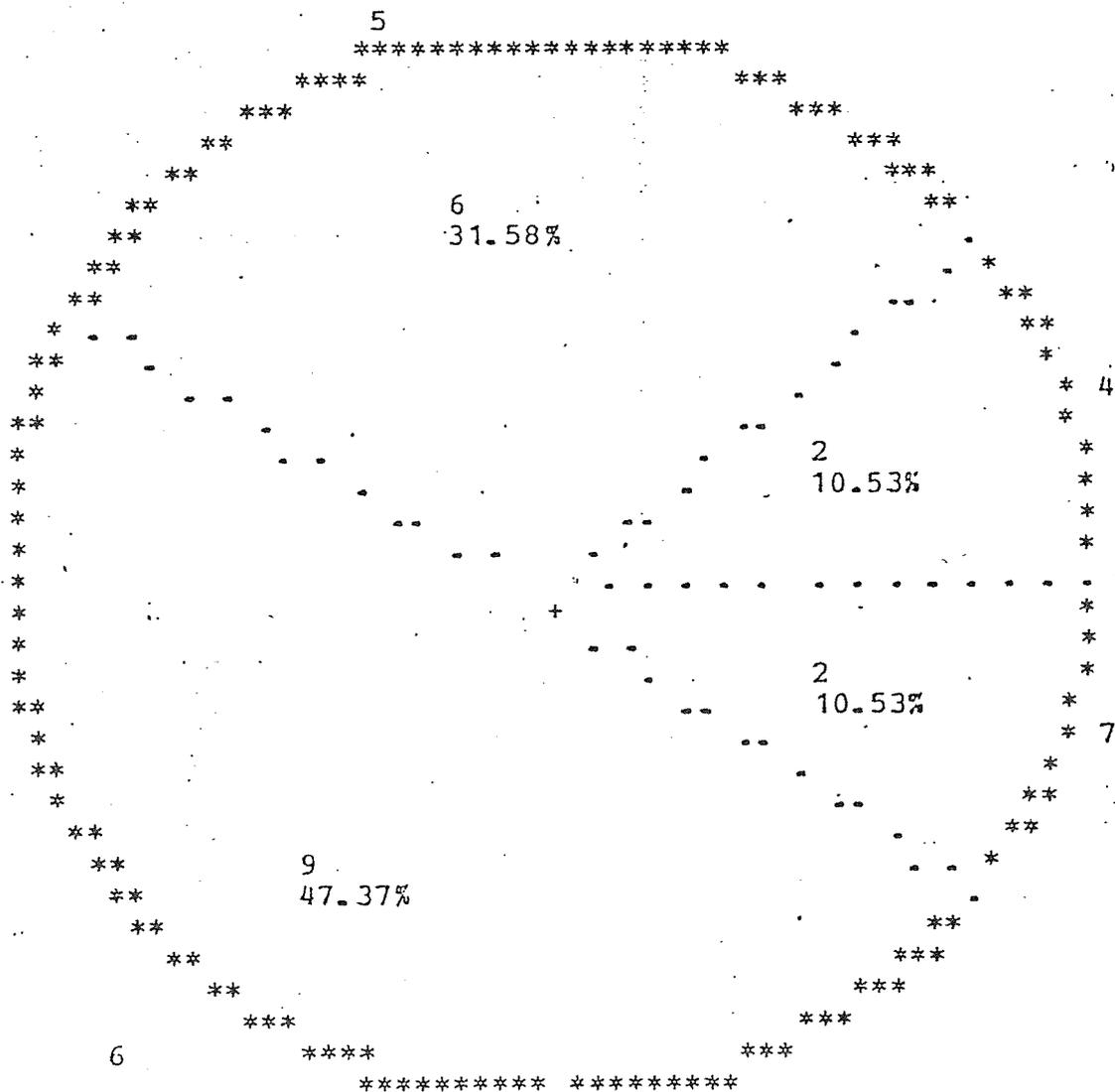


procedura chart

```
PROC CHART DATA=MEDIA.CLASSE3;  
TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';  
LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';  
PIE SCRITTO/ TYPE=REQ DISCRETE;
```

DITRIBUZIONE DEI VOTI DI ITALIANO

FREQ PIE CHART OF SCRITTO

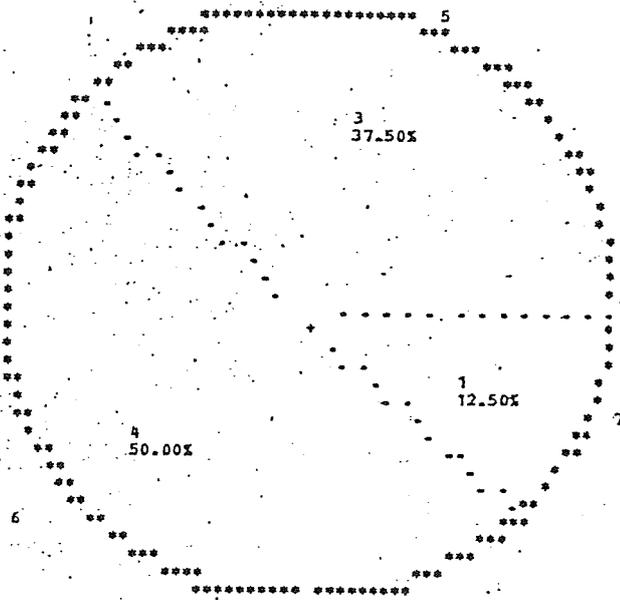


procedura chart

```
PROC SORT DATA=MEDIA.CLASSE3 OUT=SORT.CLASSE3;  
BY SSSO;  
PROC CHART DATA=SORT.CLASSE3;  
TITLE 'DISTRIBUZIONE DEI VOTI DI ITALIANO';  
LABEL SCRITTO = 'VOTO SCRITTO DI ITALIANO';  
PIE SCRITTO/TYPE=REQ DISCRETE;  
BY SSSO;
```

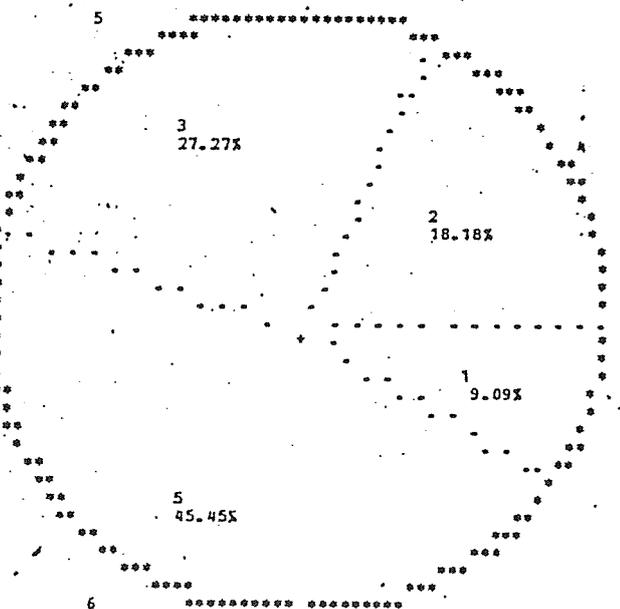
DISTRIBUZIONE DEI VOTI DI ITALIANO
SSSO=F

FREQ PIE CHART OF SCRITTO



DISTRIBUZIONE DEI VOTI DI ITALIANO
SSSO=M

FREQ PIE CHART OF SCRITTO



Dichiarazione SET

La dichiarazione SET permette la lettura di osservazioni da uno o piu' data set di tipo SAS (non si usa la INPUT !!). Nel caso si legga da un solo data set le osservazioni sono lette una per una in maniera sequenziale sino alla fine del data set. Nel caso che i data set siano due o piu' si puo' ottenere la concatenazione o l'interleaving dei data set.

Sintassi:

```
SET nome di data set SAS opzioni;  
    BY lista di variabili;
```

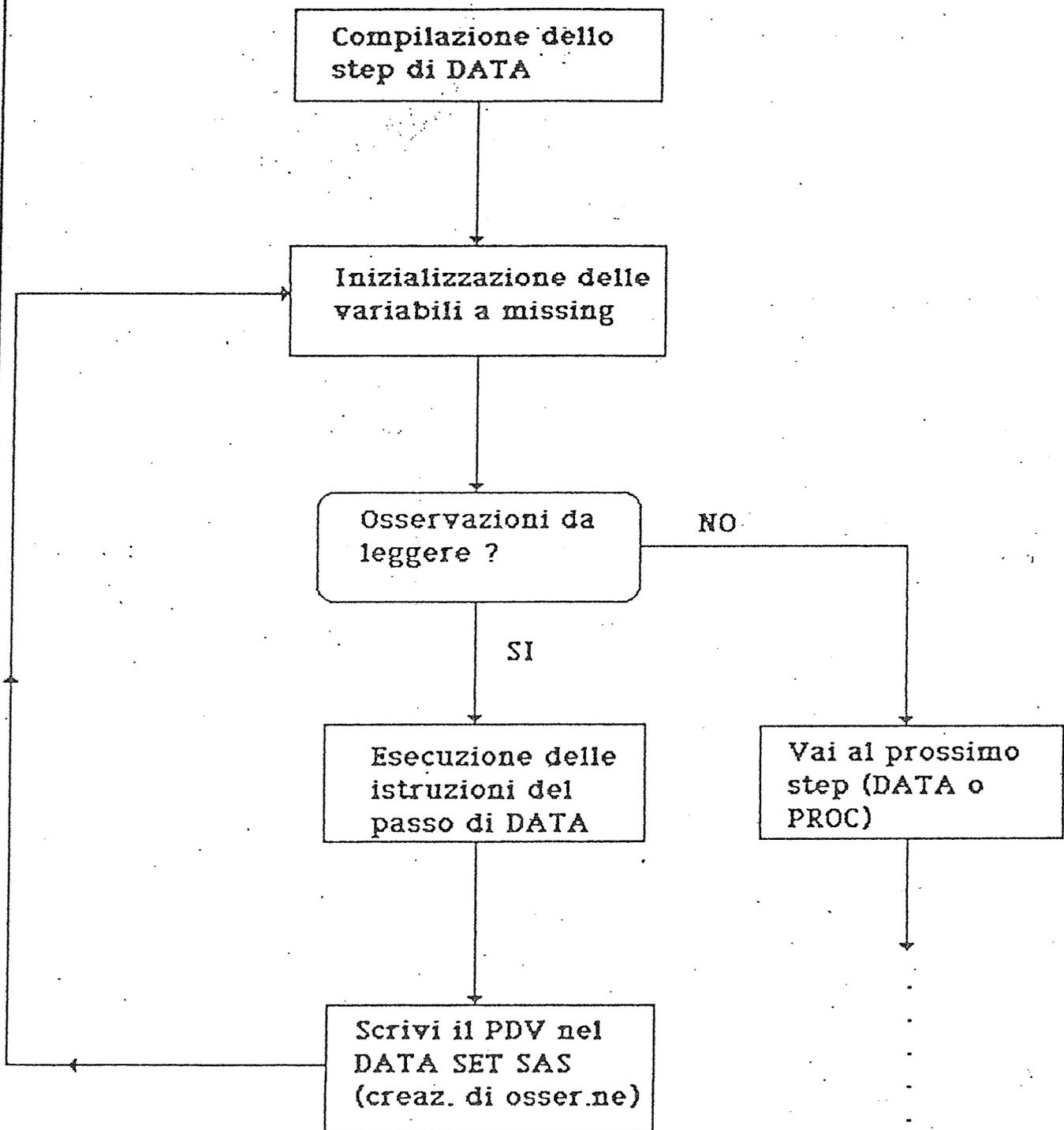
Esempi:

```
SET POPOL;  
SET FATTUR.ANN084;  
SET _LAST_;  
SET DATASET (DROP=NOME STIPEND);
```

```
SET A B; Concatena il data set B al  
data set A
```

```
PROC SORT DATA=MASCHI; BY MATRICOL;  
PROC SORT DATA=FEMMINE; BY MATRICOL;  
DATA TUTTIORD;  
    SET MASCHI FEMMINE;  
    BY MATRICOL;
```

CICLO DI LAVORO DI UN PASSO DI DATA



CAPITOLO 4

PROCEDURE STEP

S I N T A S S I

PER RICHIAMARE UNA PROCEDURA OCCORRE SCRIVERE LA SEGUENTE DICHIARAZIONE:

PROC NOME DELLA PROCEDURA OPZIONI DATA= NOME DEL DATA SET

UNO DEI NOMI VISTI NEL
PRECEDENTE EUCIDO

CIASCONA PROCEDURA
PERMETTE L'USO DI
UNA SERIE DI OPZIONI
PARTICOLARI

ESEMPIO:

```
PROC MEANS MEAN MAXDEC=3 DATA=DATASET.MIO;
```

IN MOLTI CASI, OLTRE ALLA DICHIARAZIONE PRECEDENTE OCCORRE DARE ULTERIORI INFORMAZIONI AL S.A.S., AFFINCHÈ ESSO POSSA ESEGUIRE CORRETTAMENTE LA PROCEDURA CHIAMATA.

DICHIARAZIONI INFORMATIVE

VARIABLES LISTA DI VARIABILI;

ID " " "

BY " " "

ESEMPIO: PROC SORT DATA=DATASET.DISORD OUT=DATASET.ORDIN;
 BY DATANASC;

PROC PRINT options DATA=datasetsAS;

SERVE PER STAMPARE I VALORI DELLE VARIABILI DEL DATASET

L'USO DI OPZIONI E DI STATEMENT DI CONTROLLO PERMETTE LA PERSONALIZZAZIONE DELLE STAMPE;

OPZIONI:

N STAMPA IL NUMERO DI OSSERVAZIONI DEL DATASET

DOUBLE (D) SPAZIATURA DOPPIA TRA UNA RIGA E L'ALTRA

UNIFORM CURA L'UNIFORMITA' DELLA PAGINA DI STAMPA (E' UTILE QUANDO IL DATASET HA MOLTE VARIABILI)

ROUND ARROTONDA I VALORI NUMERICI PRIMA DI TOTALIZZARLI

LABEL LE LABEL ATTRIBUITE ALLE VARIABILI SONO USATE COME INTRESTAZIONE

SPLT= DELIMITATORE INDICANTE IL PUNTO DI DIVISIONE DELLE LABEL

STATEMENTS:

BY LISTA DI VARIABILI;

SUDDIVIDE I QUADRI PER LE VARIABILI INDICATE NELLA LISTA

VAR LISTA DI VARIABILI;

LIMITA LA STAMPA ALLE SOLE VARIABILI DELLA LISTA

ID LISTA DI VARIABILI;

RIMPIAZZA LA NUMERAZIONE PROGRESSIVA CON I VALORI DELLA PRIMA VARIABILE DELLA LISTA

(ANCORA SULLA PROC PRINT)

SUM LISTA DI VARIABILI

TOTALIZZA LE VARIABILI DELLA
LISTA

LO SCOPO PRINCIPALE DELLA STATISTICA CONSISTE NEL COMPIERE UN'INFERENZA CIRCA L'INTERA POPOLAZIONE A PARTIRE DA UN CAMPIONE.

CIASCUN CAMPIONE PUO' ESSERE SEMPLIFICATO DA POCHI NUMERI DESCRITTIVI, CIASCUNO DEI QUALI SI DEFINISCE "STATISTICA CAMPIONARIA".

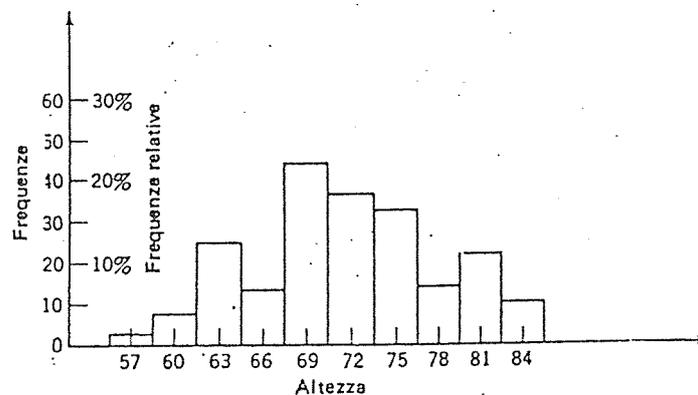
SPESSE CONVIENE RAPPRESENTARE I DATI DEL CAMPIONE IN MANIERA SINOTTICA E COMPATTA, IN MODO DA RENDERNE AGEVOLE LO STUDIO; LA TABELLA DI FREQUENZA SOTTOSTANTE NE È UN ESEMPIO:

Frequenze e frequenze relative delle stature di un campione di 200 individui.

N. delle classi	Classi limiti	Classi valore centrale	Frequenze f	Frequenze relative f/n
1	55.5-58.5	57	2	.010
2	58.5-61.5	60	7	.035
3	61.5-64.5	63	22	.110
4		66	13	.065
5		69	44	.220
6		72	36	.180
7		75	32	.160
8		78	13	.065
9		81	21	.105
10	82.5-85.5	84	10	.050

$$\sum f = 200 = n \quad \sum f/n = 1.00$$

UN ALTRO MODO DI RAPPRESENTARE GLI STESSI DATI È QUELLO DI USARE I GRAFICI DI FREQUENZA DEL TIPO MOSTRATO SOTTO:



Distribuzione delle frequenze e delle frequenze relative di un campione di 200 individui.

IL S.A.S. PERMETTE LA REALIZZAZIONE DI TABELLE DI FREQUENZA E DI GRAFICI DI FREQUENZA IN MANIERA SEMPLICE E VELOCE.

TABELLE DI FREQUENZA

LE TABELLE DI FREQUENZA POSSONO ESSERE DI **TRE** TIPI:

1) A 1 DIMENSIONE

2) " 2 DIMENSIONI

3) " 3 " "

PROC FREQ DATA=DATASET SAS;

QUESTA PROCEDURA È USATA PER GENERARE TABELLE DI FREQUENZA AD UNA O PIÙ VIE.

INOLTRE È POSSIBILE OTTENERE:

- DISTRIBUZIONE DEI VALORI DI UNA VARIABILE
- DISTRIBUZIONI CONGIUNTE DI DUE O PIÙ VARIABILI
- FREQUENZE PESATE
- MISURE DI ASSOCIAZIONE E TEST STATISTICI PER TABELLE A DUE VIE
- GENERARE DATASET CON I RISULTATI (INVECE CHE SU STAMPANTE)

STATEMENTS USABILI:

TABLES richieste di tabulazione / opzioni;

ESEMPI: TABELLA AD UNA VIA TABLES SESSO;
 TABELLA A DUE VIE TABLES SESSO*PAGA;
 TABELLA A TRE VIE TABLES CLASSE*SESSO*VOTO;

OPZIONI (ALCUNE):

EXPECTED	NOFREQ	NOCUM	ALL
DEVIATION	NOPERCENT	MISSING	LIST
CELLCHI2	NOROW	NOPRINT	SPARSE
CHISQ	NOCOL	OUT=DATASETSAS	

WEIGHT VARIABILE-PESO;

BY LISTA DI VARIABILI;

PROCEDURA CHART

LA PROCEDURA CHART E' UTILE PER DISEGNARE ISTOGRAMMI
E IDEOGRAMMI (TORTE E STELLE), SULLA STAMPANTE O SUL
TERMINALE;

SI POSSONO OTTENERE LE SEGUENTI FIGURE:

- ISTOGRAMMI ORIZZONTALI
- ISTOGRAMMI VERTICALI
- ISTOGRAMMI A BLOCCHI (TRIDIMENSIONALI)
- TORTE (PIE)
- STELLE

FORMATI

Cos'è un formato? PUO' ESSERE ASSIMILATO AD UN "FILTRO" CHE DEFINISCE "COME" IL VALORE DI UNA VARIABILE DEVE ESSERE PRODOTTO IN USCITA/ENTRATA

ESISTONO FORMATI GIÀ "PRECONFEZIONATI" NEL SAS ED IL LORO SPETTRO DI APPLICAZIONE È MOLTO AMPIO:

ESEMPI: -W. W.D SW. COMMAW.D
 ROMANW. WORDSW. BESTW. ECC. ECC..

L'ASSOCIAZIONE DI UNA FORMATO AD UNA O PIÙ VARIABILI AVVIENE TRAMITE LO STATEMENT FORMAT:

ESEMPIO: FORMAT ALTEZZA PESO 3. ;
 FORMAT NOME S30. ALTEZZA 3. PESO 2.;

N.B.: SE QUESTO STATEMENT VIENE USATO IN UN PASSO DI DATA, ESSO VIENE ASSOCIATO ALLA VARIABILE IN MANIERA PERMANENTE SINO AD UNA NUOVA ASSOCIAZIONE TRAMITE FORMAT. SE VIENE USATA IN UNA PROC, IL SUO EFFETTO SARÀ VALIDO PER LA SOLA DURATA DELLA PROC.

CREAZIONE DI FORMATI "AD HOC"

E' POSSIBILE CREARE UN FORMATO PER I PROPRI SCOPI USANDO LA PROCEDURA:

```
PROC FORMAT opzioni;
```

OPZIONI:

```
PRINT DDNAME=ddname.
```

STATEMENTS USABILI:

```
VALUE nome del formato (opzioni)
```

```
range1=label1
```

```
range2=label2
```

```
.....  
.....;
```

```
PICTURE nome del formato (opzioni)
```

```
range1='picture1' (opzioni)
```

```
range2='picture2' (opzioni)
```

```
.....  
.....;
```

NOTE: INFILE COMPON IS FILE COMPON FILE A1
NOTE: 7 LINES WERE READ FROM INFILE COMPON.
THE MINIMUM LINE LENGTH IS 120.

THE MAXIMUM LINE LENGTH IS 120.

NOTE: DATA SET WORK,GIUNTA HAS 7 OBSERVATIONS AND 20 VARIABLES.
NOTE: THE DATA STATEMENT USED 0.12 SECONDS AND 832K.

```

28      PROC FORMAT;
29      VALUE LISTFORM 01 = PCI
30      02 = PSI
31      03 = PDUP
32      04 = DP
33      05 = 'IND-SINISTRA'
34      06 = PSDI
35      07 = PRI
36      08 = DC
37      09 = PLI
38      10 = 'IND-CENTRO'
39      11 = 'MSI-DN'
40      12 = 'ALTRI-LISTA LOCALE'
41      13 = 'PART. RADICALE';

```

NOTE: THE PROCEDURE FORMAT USED 0.10 SECONDS AND 832K.

CAPITOLO 5

ESEMPI DI PROGRAMMI

data step

DATA MEDIA.CLASSE3;
INPUT NOME \$25.
 SESSO \$ 27-27
 +2 ETA 2.
 @35 SCRITTO 35.
 @40 ORALE 1.;

CARDS;

KORI GINO	M	14	5	5
SORRI DANILO	M	13	7	6
LIPERINI FABRIZIO	M	15	.	5
TOZZINI LOREDANA	F	14	5	6
BERTINI MARIA	F	14	6	6
GARZELLI PAOLO	M	14	.	.
GHELARDI ANNA	F	14	6	7
ANIA GIUSEPPE	M	15	4	4
MICHELETTI BRUNO	M	14	6	5
GARZELLI VITTORIO	M	14	6	6
BIGOTTI MIRIANA	F	14	5	6
TONINI RENZO	M	16	4	5
FANTONI CINZIA	F	13	6	6
BALLONI FRANCA	F	14	7	8
ACERBI NILO	M	14	6	5
GINI ALDO	M	15	6	6.
VALLINI CLAUDIA	F	14	5	5
CORONA GIOVANNI	M	14	6	5
MATTEI DINO	M	14	5	5
BELLANDI ANGELA	F	14	6	6
MARRAZZO UMBERTO	M	14	5	6

;

procedura contents

PROC CONTENTS DATA=MEDIA.CLASSE3;

CONTENTS OF SAS DATA SET MEDIA.CLASSE3

OBSERVATIONS=21 CREATED BY CMS USERID MESURE ON CPUID FF-3081-001964

AT 15:00 TUESDAY, APRIL 30, 1985 BY SAS RELEASE 82.3

FILE=MEDIA CLASSE3 BLKSIZE=8158 LRECL=54 GENERATED BY DATA

ALPHABETIC LIST OF VARIABLES

#	VARIABLE	TYPE	LENGTH	POSITION	FORMAT	INFORMAT
3	ETA	NUM	8	30		
1	NOME	CHAR	25	4		
5	ORALE	NUM	8	46		
4	SCRITTO	NUM	8	38		
2	SESSO	CHAR	1	29		

----- SOURCE STATEMENTS -----

```
| DATA MEDIA.CLASSE3;  
| INPUT NOME $25.  
|     SESSO $ 27-27  
|     +2 ETA 2.  
|     @35 SCRITTO 35  
|     @40 ORALE 1.;  
| CARDS;
```

procedura contents.

```
PROC CONTENTS DATA=MEDIA ALL NOSOURCE;
```

SAS DATA SET DIRECTORY

NAME	# OBS
CLASSE2	21
CLASSE1	21
CLASSE3	21

CONTENTS OF SAS DATA SET MEDIA.CLASSE1

OBSERVATIONS=21 CREATED BY CMS USERID MESURE ON CPUID PP-3081-001964
AT 14:58 TUESDAY, APRIL 30, 1985 BY SAS RELEASE 82.3

FILE=MEDIA CLASSE2 BLKSIZE=8158 LRECL=54 GENERATED BY DATA

ALPHABETIC LIST OF VARIABLES

#	VARIABLE LABEL	TYPE	LENGTH	POSITION	FORMAT	INFORMAT
3	ETA	NUM	8	30		
1	NOME	CHAR	25	4		
5	ORALE	NUM	8	46		
4	SCRITTO	NUM	8	38		
2	SESSO	CHAR	1	29		

CONTENTS OF SAS DATA SET MEDIA.CLASSE2

OBSERVATIONS=21 CREATED BY CMS USERID MESURE ON CPUID PP-3081-001964
AT 14:58 TUESDAY, APRIL 30, 1985 BY SAS RELEASE 82.3

FILE=MEDIA CLASSE1 BLKSIZE=8158 LRECL=54 GENERATED BY DATA

ALPHABETIC LIST OF VARIABLES

#	VARIABLE LABEL	TYPE	LENGTH	POSITION	FORMAT	INFORMAT
3	ETA	NUM	8	30		
1	NOME	CHAR	25	4		
5	ORALE	NUM	8	46		
4	SCRITTO	NUM	8	38		
2	SESSO	CHAR	1	29		

procedura contents

```
PROC CONTENTS DATA=MEDIA.CLASSE3 SHORT;
```

CONTENTS OF SAS DATA SET MEDIA.CLASSE3

ALPHABETIC LIST OF VARIABLES

ETA NOME ORALE SCRITTO SESSO

procedura sort

```
PROC SORT DATA=MEDIA.CLASSE3;  
  BY NOME;  
PROC PRINT DATA=MEDIA.CLASSE3;
```

OBS	NOME	SESSO	ETA	SCRITTO	ORALE
1	ACERBI NILO	M	14	6	5
2	ANIA GIUSEPPE	M	15	4	4
3	BALLONI FRANCA	F	14	7	8
4	BELLANDI ANGELA	F	14	6	6
5	BERTINI MARIA	F	14	6	6
6	BIGOTTI MIRIANA	F	14	5	6
7	CORONA GIOVANNI	M	14	6	5
8	FANTONI CINZIA	F	13	6	6
9	GARZELLI PAOLO	M	14	.	.
10	GARZELLI VITTORIO	M	14	6	6
11	GHELARDI ANNA	F	14	6	7
12	GINI ALDO	M	15	6	6
13	LIPERINI FABRIZIO	M	15	.	5
14	MARRAZZO UMBERTO	M	14	5	6
15	MATTEI DINO	M	14	5	5
16	MICHELETTI BRUNO	M	14	6	5
17	HORI GINO	M	14	5	5
18	SORRI DANILO	M	13	7	6
19	TONINI RENZO	M	16	4	5
20	TOZZINI LOREDANA	F	14	5	6
21	VALLINI CLAUDIA	F	14	5	5

