**SESSIÓ 3 – LECTURA DE FITXERS**

\*LECTURA DATOS EN FICHEROS DE TEXTO, PLANOS O ASCII;

/\*COMANDO INPUT

1 INPUT nv1 nv2 .... nvk $ .....;

2 INPUT nv1 p1-p2 nvk $ pk-pk2 ....; pi -->num col donde se encuentra el valor a leer

3 INPUT @p1 V1 informat1w. .........@pk Vk informatkw.;

3 obligatorio si teneis informació no estandard (información estandar: son varaibles numericas -sep decim . y cadenas)

Información no estándar: dates, $123, 123,152,527 ...

4 INPUT como un mix de tot lo anterior\*/

\*1 INPUT

Los datos son de tipo estandard,

si hay varaibles texto deben ser cortas (max 8 cars)

No existen datos missng representados por espacios en blanco

Los valores de las var deben estar separados entre si por un o más blancos;

**DATA** EX\_INPUT1;

INFILE 'ruta\nom.ext' opcions ; /\*opcions de infile\*/

INPUT V1 V2 V3$ ;

**RUN**;

**DATA** EX\_INPUT1;

INPUT V1 V2 V3$ ;

\*más comandos;

datalines;

1 3 h

5 6 d

4 8 d

;

\*EJ 1 Crearemos la base S3\BASES\_SAS\bas1.sas7bdat que tendrá

2 variables: X1( nombre) y X2 numérica;

\*LIBNAME --> S3\BASES\_SAS;

libname d 'd:\s3\bases\_sas';

**DATA** d.EX1;

INPUT nom$ **1**-**12** genere;

DATALINES;

PEDRO 1

ELENA 0

MARIO 1

JOAN MARCOS 1

;

**PROC** **PRINT** data=d.EX1;

**RUN**;

/\*EJ 2 Examinar el archivo de datos ALCADES.DAT (altura a principio y a final de año).

Escribir un programa para leer los datos y guardarlos temporalmente en alt.sas7bdat,

hacer una descriptiva inicial de las variables ,

y crear variable altura promedio y describirla\*/

**DATA** alt;

infile 'D:\S3\Bases\_sas\ALCADES.DAT' ;

input alti altf;

if (alti=**0** | alti=-**1**) then alti=**.**;

altm=(alti+altf)/**2**;

**run**;

**PROC** **MEANS** DATA = ALT; **RUN**;

/\* 2 INPUT var1 p11-p1j var2 p21-p2j . vark pk1-pkj;

Dónde:

VARK Simboliza el nombre de la variable k-ésima

PK1 --> número columna en la que debe situarse el puntero de lectura para iniciar

la lectura del valor de la VARK

PKj --> número columna en la que debe situarse el puntero de lectura para finalizar

la lectura del valor de la VARK

IMPORTANTE --> Los datos a leer deben estar debidamente alineados en columnas.

Características de los datos para poder utilizar COLUMN INPUT.

a. La información debe ser lo que SAS llama de tipo estándar:

números sin ningún carácter especial salvo el punto como separador de decimales

cadena o alfanumérica

b. Los valores de las variables no están necesariamente separados por espacios en blanco.

c Las variables pueden presentar valores missing representados por espacios en blanco

d. Podemos leer sólo la información relativa a algunas variables.

e. Las variables cadena pueden tener una longitud superior a 8 caracteres y/o blancos en su interior.

3.2.1 Column input (la información debe ser estándar)\*/

\*EJ 3 Crearemos la base S3\BASES\_SAS\EX3.SAS7BDAT

con 3 variables edad, peso, genero \_caracter

utilizando Column input ;

**DATA** d.ex3;

input edad peso genero $ **10**;

datalines;

23 87.70 H

28 54 D

21 D

32 59 H

54 62 D

;

**proc** **print** data=d.ex3;

**run**;

\*EJ 4

a) Crear la bdd S3\BASES\_SAS\EX4.SAS7BDAT a partir del fichero S3\EJERCICIOS\NOTES.TXT

pero sólo con las variables NOMBRE, NOTA\_R Y NOTA\_SAS

b)Crear la variable NOTAF como promedio de las dos notas;

**DATA** d.EX4;

infile 'D:\S3\Bases\_sas\NOTES.TXT';

input nombre$ nota\_r **13**-**15** nota\_sas **17**-**19**;

nota=(nota\_r+nota\_sas)/**2**;

**run**;

**proc** **print** data = d.ex4; **run**;

\*EJ 5

a) Crear la BDD S3\BASES\_SAS\EX5.SAS7BDATa partir del fichero S3\EJERCICIOS\NOTES\_M.TXT

con las variables NOM, SEXE, EDAD, NOTA\_R Y NOTA\_SAS

b) Crear la variable NOTAF como promedio de las dos notas;

**DATA** d.EX5;

infile 'D:\S3\Bases\_sas\notes\_m.txt' TRUNCOVER;

INPUT nombre$ sexe$ edat nota\_r **13**-**15** nota\_sas ;

nota=(nota\_r+nota\_sas)/**2**;

**run**;

**proc** **print** data = d.ex5; **run**;

\*EJ6;

/\*leer la información de los archivos

dadesnot.dat y pulsesas.dat

la información de estos dos ficheros está en ASCII

DADESNOT.DAT

Contiene:

genero 1 = hombre 2 = mujer

erasmus si/no

nota1 sobre 100

nota2 sobre 10

A partir del dadesnot.dat

- crear la base temporal notes.sas7bdat

con todas las variables y todas las observacions\*/

\* Hacer un descriptivo breve de las variables

para detectar posibles anomalias en los datos (valores fuera de rango)

en caso q encontreis errores efectuar la correción;

**DATA** NOTES;

infile 'D:\S3\Bases\_sas\dadesnot.DAT' truncover;

input genero **1** erasmus$ **2**-**3** nota1 **4**-**5** nota2;

**RUN**;

**proc** **print** data=notes;**run**;

**proc** **means** data=notes;**run**;

**data** notes;

set notes;

if nota2=**99** then nota2=**.**;

if nota1=-**1** then nota1=**.**;

;

**proc** **means** data=notes;**run**;

\*EJ7;

/\*leer la información de pulsesas.dat

solo queremos las 2 primeras variables numéricas

nom p1 y p2 \*/

**DATA** PULSA (keep= p1 p2);

infile 'D:\S3\Bases\_SAS\PULSESAS.DAT';

input nom$ p1 p2 v1 v2 v3 v4 v5 v6;

LABEL P1 = Pulsaciones antes del esfuerzo

P2 = Pulsaciones después del esfuerzo;

**RUN**;

**proc** **print** data=pulsa L;**run**;

**proc** **means** data=pulsa;**run**;

\* Hacer un descriptivo breve de las variables

para detectar posibles anomalias en los datos (valores fuera de rango)

en caso q encontreis errores efectuar la correción;