



Noções de SAS®

Maria Teresa Leão Costa

2023

SAS

SAS

O SAS é um sistema de aplicação integrada, que consiste em vários produtos que tem por função: acesso, gerenciamento, análise estatística e apresentação de dados, somada a uma linguagem poderosa de programação e geração de relatórios.

■ **Acesso aos dados** - o sistema SAS possui ferramentas para acessar os bancos de dados mais populares, tratando suas tabelas como arquivos SAS.

■ **Gerenciamento de dados** - com o SAS é possível editar, selecionar, ordenar e juntar arquivos.

(3)

SAS

■ **Análise de dados** - o SAS possui rotinas estatísticas para análises de regressão, análise de variância, de componentes principais, discriminante, modelos lineares, análise fatorial, previsões, controle de qualidade, cálculos matriciais, etc.

■ **Apresentação de informações** - a geração de relatórios é fácil e flexível. Eles podem ser listados, tabulados e gráficos.

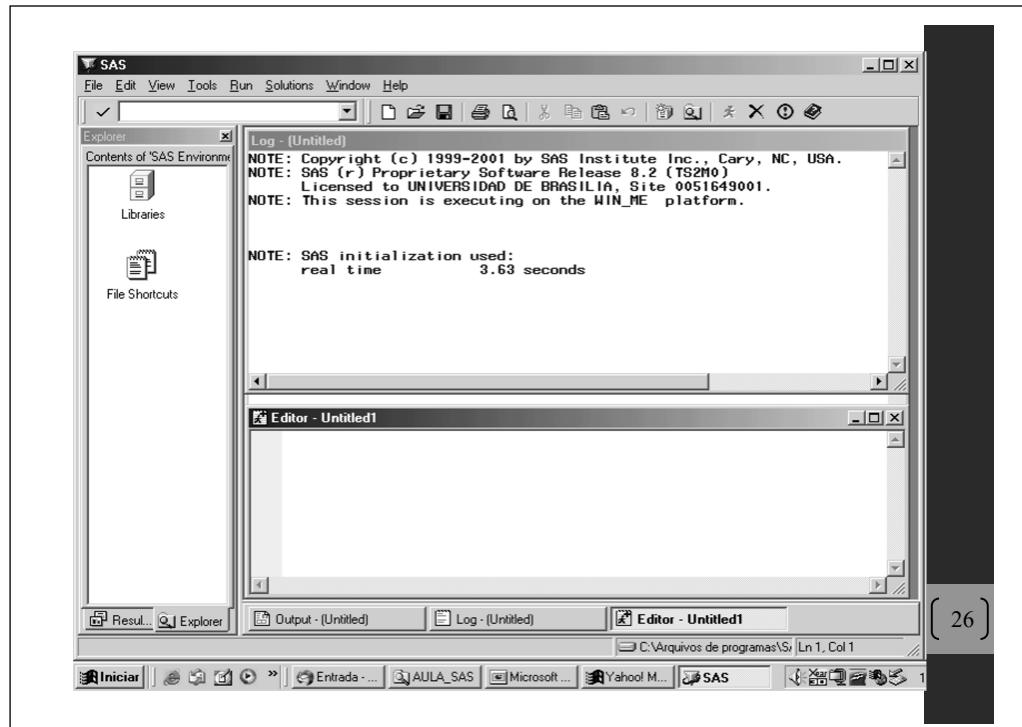
(4)

ÁREAS DE TRABALHO DO SAS

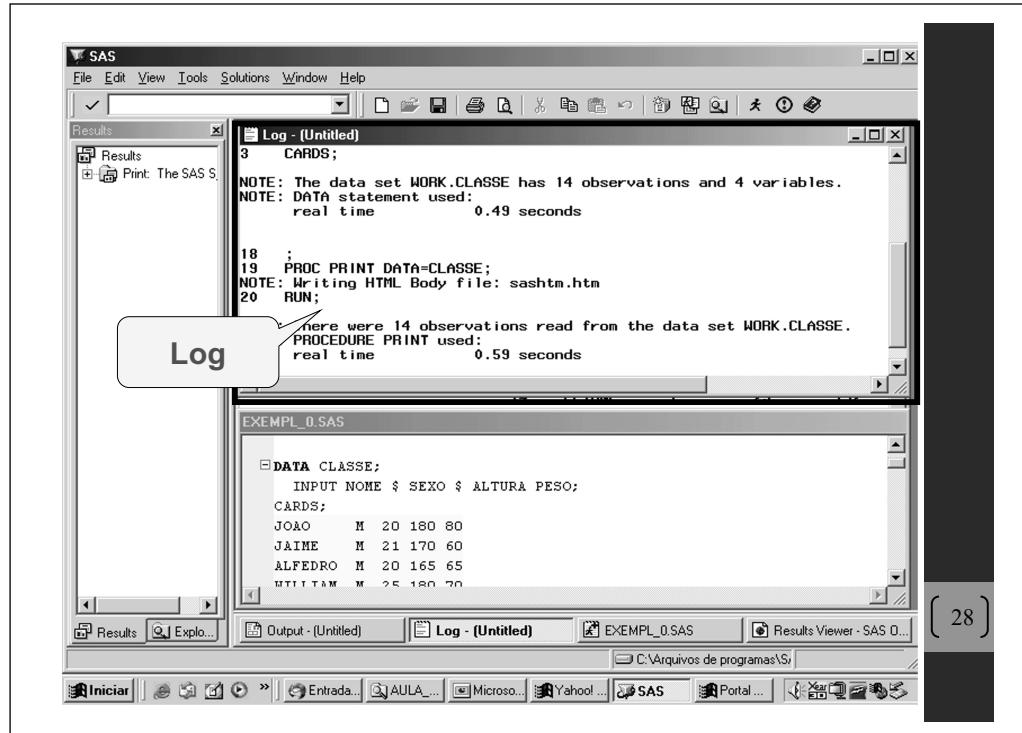
A área de trabalho do SAS é composta de janelas que permitem realizar tarefas específicas.

- Janelas principais de programação:
 - Program Editor
 - Log
 - Output
 - Explorer
 - Results

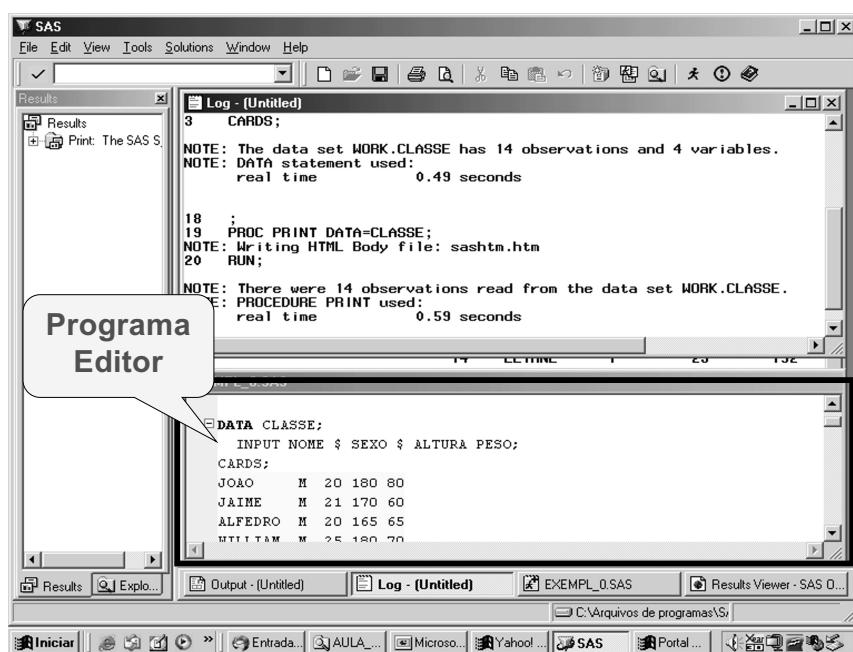
[25]



[26]



[28]



[27]

Output

Obs	NOME	SEXO	IDADE	ALTURA	PESO
1	JOAO	CARLOS	21	170	80
2	JAIME	M	21	170	60
3	ANA	LUCIA	.	23	140
4	ROBERTO	M	25	165	55
5	JANETE	F	20	173	60
6	LUIZO	M	19	170	65
7	LUIZA	F	20	165	55
8	MARIA	F	24	150	50
9	LORENA	F	25	160	50
10	ANGELA	F	20	160	52

[29]

Explorer

```

NOTE: Copyright (c) 2002-2003 by SAS Institute Inc., Cary, NC, USA.
NOTE: SAS (r) 9.1 (TS1M3)
      Licensed to UNIVERSIDAD DE BRASILIA, Site 0051649001.
NOTE: This session is executing on the XP_HOME platform.

NOTE: 0.49 seconds
      0.49 seconds

Log - (Untitled)

/* PRIMEIRO PROGRAMA */

DATA CLASSE;
  INPUT NOME $ SEXO $ IDADE ALTURA PESO;
  CARDS;
  JOAO CARLOS M 20 180 80
  JAIME M 21 170 60
  ANA LUCIA F 23 140 46
  ROBERTO M 25 165 55
  JANETE F 20 173 60

```

[30]

Results

```

3  CARDS;
NOTE: The data set WORK.CLASSE has 14 observations and 4 variables.
NOTE: DATA statement used:          0.49 seconds
18
19
20 RUN;

NOTE: There were 14 observations read from the data set WORK.CLASSE.
NOTE: PROCEDURE PRINT used:        real time 0.59 seconds

EXEMPL_0.SAS

DATA CLASSE;
  INPUT NOME $ SEXO $ ALTURA PESO;
  CARDS;
  JOAO   M 20 180 80
  JAIME  M 21 170 60
  ALFEDRO M 20 165 65
  MUITITAM M 25 180 70

```

[31]

Como Trabalhar com o SAS® ?

Opções de uso do SAS para análise de dados:

- Tradicional - programando
- Enterprise Guide
- ...

[32]

CONHECENDO O PROGRAMA SAS

```
DATA CLASSE;
  INPUT NOME $ SEXO $ ALTURA PESO;
CARDS;
JOAO      M  20 180 80
JATIME    M  21 170 60
ALFEDRO   M  20 165 65
WILLIAM   M  25 180 70
ANA        F  23 140 46
ALICE      F  24 155 50
ROBERTO   M  25 165 55
LORENA     F  25 160 50
ANGELA     F  20 160 52
ELIANE     F  23 152 52
;
PROC SORT DATA=CLASSE OUT=CLASSE1;
  BY NOME;
RUN;

PROC PRINT DATA=CLASSE1;VAR NOME SEXO; RUN
;
```

[33]

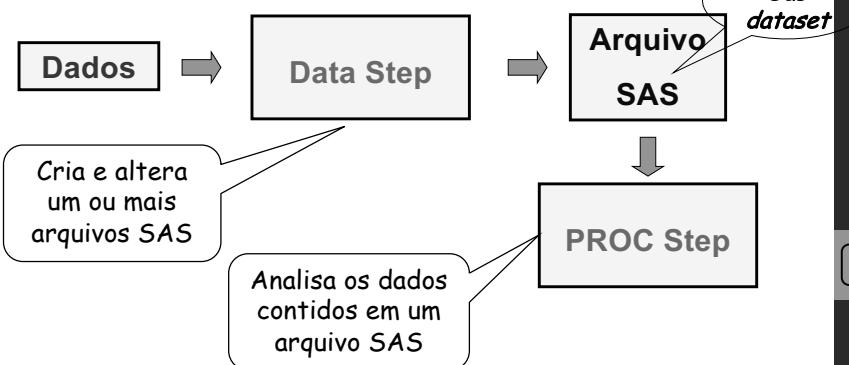
Regras para escrever comandos SAS

- Os comandos começam com uma palavra chave e terminam com um ponto e vírgula.
- Os comandos SAS são formatados livremente:
 - eles podem começar em qualquer parte e terminar em qualquer parte;
 - um comando pode continuar em outras linhas;
 - vários comandos podem estar em uma mesma linha.
- Recomenda-se que os comandos DATA e PROC comecem na 1^a coluna de seu PROGRAM EDITOR.

[34]

Estrutura de um Programa SAS

- Todos os dados deverão ser inseridos em arquivos com estrutura do SAS (**DATA SET**), para serem analisados pelos procedimentos do SAS (**PROC**).



[35]

[36]

```

DATA CLASSE;
  INPUT NOME $ SEXO $ ALTURA PESO;
CARDS;
JOAO      M  20 180 80
JAIME     M  21 170 60
ALFEDRO   M  20 165 65
WILLIAM   M  25 180 70
ANA        F  23 140 46
ALICE      F  24 155 50
ROBERTO   M  25 165 55
LORENA     F  25 160 50
ANGELA     F  20 160 52
ELIANE     F  23 152 52
;
PROC SORT DATA=CLASSE OUT=CLASSE1;
  BY NOME;
RUN;
PROC PRINT DATA=CLASSE1;VAR NOME SEXO; RUN
;

```

DATA
Step

[37]

PROC
Step

ENTRADA DE DADOS

[38]

Criando Arquivos SAS



- Um arquivo SAS pode ser temporário ou permanente.
- Em um único programa SAS, podem ser criados e analisados vários arquivos.
- Todo arquivo SAS tem nome que:
 - Possui de 1 a 32 caracteres;
 - Começa com letra (A-Z);
 - Pode continuar com números letras ou travessões.

[39]

Entrada de Dados

- Via programa SAS:

```

DATA nome do arquivo ;
  INPUT lista de variáveis ;
  DATALINES;          (ou CARDS;)
                    { dados }
                    {
  ;

```

[40]

```

/*
***** ENTRADA DE DADOS *****
*/
DATA CLASSE;
  INPUT NOME $ SEXO $ ALTURA PESO;
CARDS;
JOAO    M  20 180 80
JAIME   M  21 170 60
ALFEDRO M  20 165 65
WILLIAM M  25 180 70
ANA     F  23 140 46
ALICE   F  24 155 50
ROBERTO M  25 165 55
JANETE  F  20 173 60
LUCIO   M  19 170 65
LUIZA   F  20 165 55
MARIA   F  24 150 50
LORENA  F  25 160 50
ANGELA  F  20 160 52
ELIANE  F  23 152 52
;

```

[41]

Entrada de Dados

- **Dados Externos (HD, CD ou pendrive):**

```

DATA nome do arquivo ;
INFILE 'diretório:\ nome do arquivo externo.extensão'
;
INPUT lista de variáveis ;

```

```

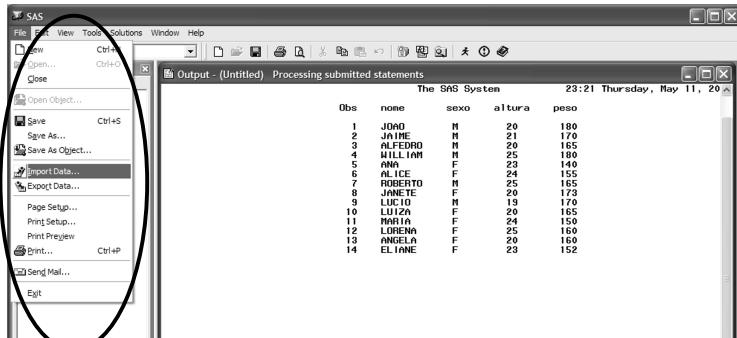
data classe1;
  infile "e:\aula_sas\dados.txt";
  input nome $ sexo $ idade altura peso;

```

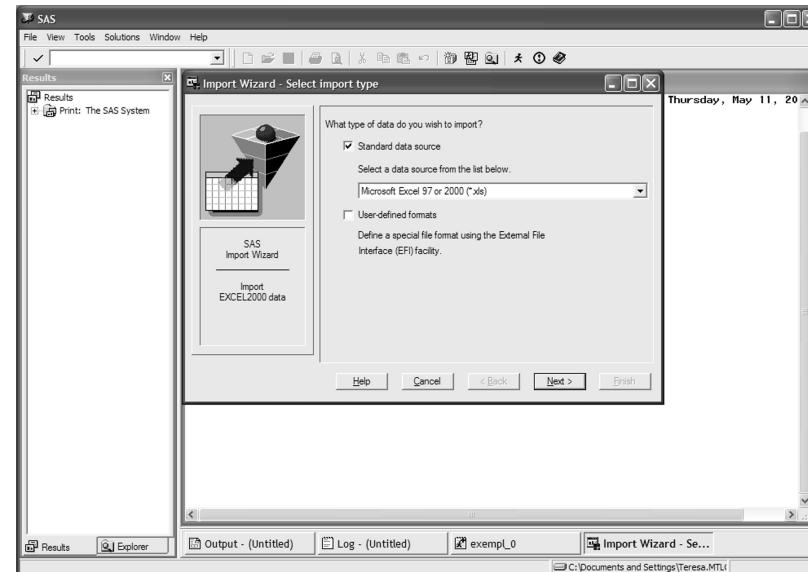
[42]

Entrada de Dados

- **Via importação de dados:**



[43]



[44]

DATA CLASSE;

INPUT NOME \$ SEXO \$ ALTURA PESO;

CARDS;

JOAO M 20 180 80

JAIME M 21 170 60

ALFEDRO M 20 165 65

WILLIAM M 25 180 70

ANA F 23 140 46

ALICE F 24 155 50

ROBERTO M 25 165 55

JANETE F 20 173 60

LUCIO M 19 170 65

LUIZA F 20 165 55

MARIA F 24 150 50

LORENA F 25 160 50

ANGELA F 20 160 52

ELIANE F 23 152 52

;

[41]

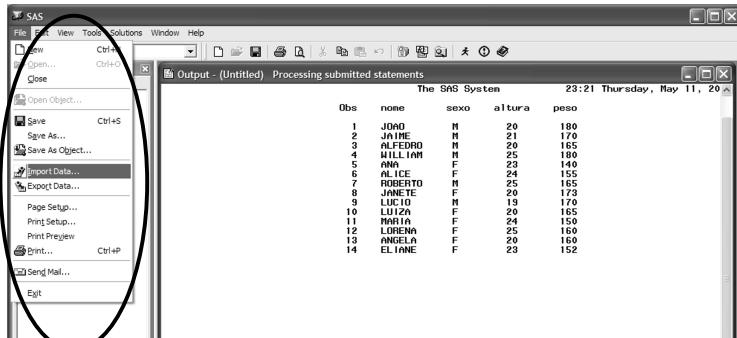
data classe1;

infile "e:\aula_sas\dados.txt";

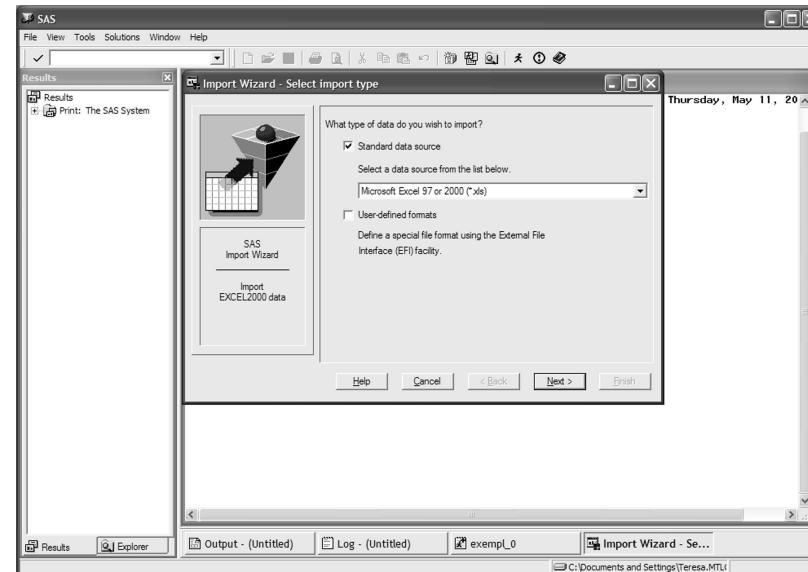
input nome \$ sexo \$ idade altura peso;

[42]

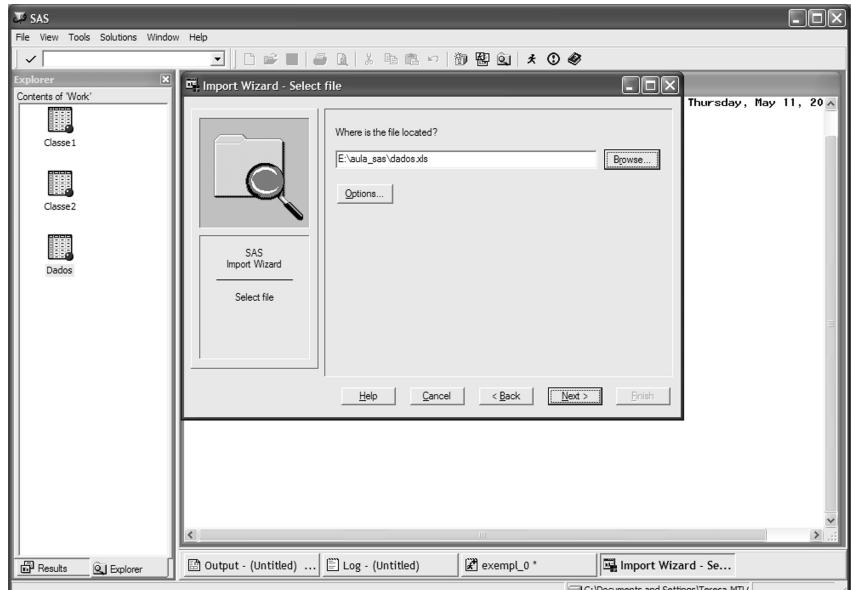
- **Via importação de dados:**



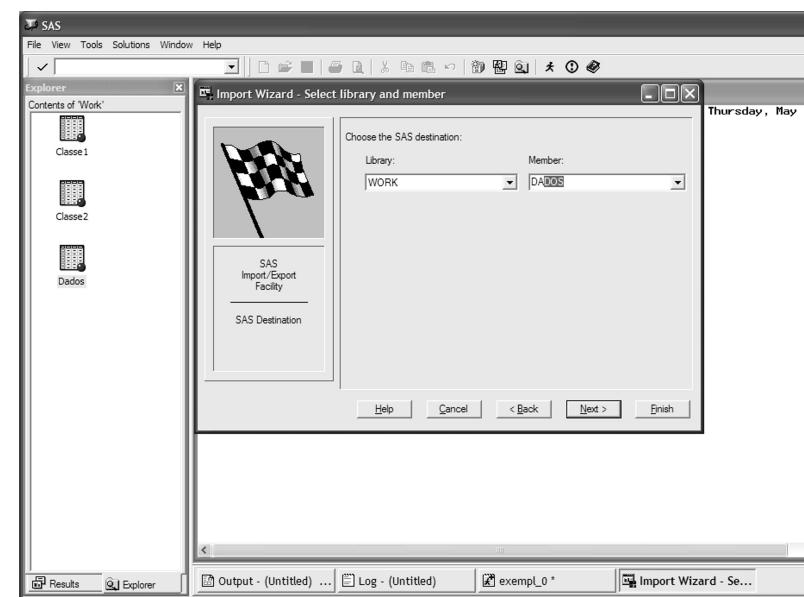
[43]



[44]



[45]



[46]

- Via PROC IMPORT:
- Arquivo Excel:

PROC IMPORT

```
DATAFILE= 'diretório:\ nome do arquivo.xlsx'
OUT= nome do arquivoSAS DBMS= xlsx REPLACE ;
GETNAMES= YES ;
RUN ;
```

[47]

Comandos do DATA STEP

Comando DATA:

- Assinala o início da DATA STEP.
- Dá nome ao arquivo a ser criado no DATA STEP.
- Para ler dados dentro de um arquivo SAS é preciso:
 - Definir a fonte de dados (comandos DATALINES ou INFILE)
 - Definir a localização das variáveis na fonte de dados (comando INPUT)

[49]

Comando **CARDS** ou **DATALINES**:

- Indica que os dados estão a seguir no programa.

Comando **INFILE**:

- Identifica o arquivo externo de dados (somente para leitura) do qual os dados serão lidos.

[50]

• **INPUT LISTADO**

- Indica as variáveis a serem lidas no arquivo.
- As mesmas são lidas na ordem em que aparecem no **INPUT**.

• **SINTAXE:**

```
INPUT variável1 [$] variável2 [$] ... Variávelk [$];
```

• **Exemplo:**

```
DATA CLASSE;  
INPUT NOME $ SEXO $ ALTURA PESO;
```

[52]

Comando **INPUT**:

- Define o nome, localização, e tipos de cada variável de um arquivo

• Tipos básicos de comando INPUT:

- INPUT LISTADO
- INPUT COLUNADO
- INPUT FORMATADO

[51]

• **REGRAS E RESTRIÇÕES:**

- O sinal **\$** é usado para indicar variáveis alfanuméricas;
- Os campos devem ser especificados na ordem;
- Os campos devem ser separados por brancos
- Use “.” como indicador de *missing value*.

• **QUANDO USAR:**

- Arquivos onde os dados são armazenados com um espaço entre cada campo, isto é, sem campos pré-fixados.
- Usado freqüentemente para entrada de dados via programa.

[53]

- **INPUT COLUNADO**

- Especifica as variáveis a serem lidas no arquivo, indicando a localização dos campos o registro de leitura de dados pela posição da coluna.
- Permite também especificar o sinal e o número de casas decimais.

- **SINTAXE:**

```
INPUT variável1 [$] col início-col fim [.decimal] variável2 [$] col início-col fim [.decimal] ..... variávelk [$] col início-col fim [.decimal];
```

- **EXEMPLO:**

```
DATA CLASSE;
  INPUT NOME $ 1-11 INICIAL $ 1 SEXO $ 13 IDADE 16-17
        ALTURA 19-21 .2 PESO 23-24;
DATALINES;
```

(54)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C	A	R	L	O	S			M	1	8		1	7	0	6	8	.	0	
C	A	R	L	A				F	2	6		1	7	8	7	2	.	0	

```
DATA FICHA ;
  INPUT NOME $ 1-8 SEXO $ 9 IDADE 11-12 ALTURA 14-16 .2 PESO 17-20 ;
CARDS ;
<dados da tabela acima>
;
```

(55)

- **REGRAS E RESTRIÇÕES:**

- O sinal \$ é usado para indicar variáveis alfanuméricas;
- As posições das variáveis devem ser pré-fixadas;
- Os campos com os dados podem ser lidos em qualquer ordem;
- Campos em branco (sem informação) são considerados como *missing values*.
- É permitido “brancos” em campos tipo alfanumérico.
- Campos ou partes de campos podem ser reidos.

(56)

- **QUANDO USAR:**

- Arquivos onde os dados são armazenados com um espaço entre cada campo, isto é, sem campos pré-fixados.
- Usado freqüentemente para entrada de dados via programa.

- **INPUT FORMATADO**

- Lê as variáveis especificando o tamanho e o formato (opcional) do campo da linha de dados.

- **SINTAXE:**

```
INPUT variável [$] w.d ou formato w.d;
```

- **EXEMPLO:**

```
DATA UM;
  INPUT ALTURA 4.2 PESO 4. ;
DATALINES;
```

57

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
C	A	R	L	O	S			M	1	8		1	7	0	6	8	.	0	
C	A	R	L	A			F	2	6		1	7	8	7	2	.	0		

DATA FICHA :
 INPUT NOME \$7. SEXO \$1. IDADE 3.
 ALTURA 4.2 PESO 4.;

CARDS :
 <dados da tabela acima>
 ;

[58]

Introdução ao PROC STEP

[59]

Introdução ao PROC STEP

A PROC é um comando utilizado para chamar um procedimento SAS e iniciar uma PROC STEP.

Os procedimentos SAS são rotinas que:

- Lêem arquivos SAS;
- Computam estatísticas;
- Imprimem resultados;
- Geram gráficos;
- Criam novos arquivos SAS;

[60]

SINTAXE:

```
PROC nome do procedimento opções ;
  comando ;
  comando ;
RUN ;
```

nome do procedimento – palavra-chave que identifica o procedimento a ser executado.

opções – parâmetros opcionais de configuração para execução do procedimento.

comando – comandos de detalhamento na execução do procedimento.

[61]

Comandos Comuns das PROC's

Comando VAR

- Especifica as variáveis do arquivo SAS que serão processadas pelo procedimento.

```
VAR <variáveis> ;
```

Comando BY

- Especifica uma ou mais variáveis que permitirão o processamento de subgrupos.

```
BY <variáveis> ;
```

[62]

Comando TITLE

- Especifica um ou mais títulos para o relatório a ser gerado pelo procedimento.

```
TITLE "<cabeçalho>" ;
```

Comando FOOTNOTE

- Especifica um ou mais rodapés para o relatório a ser gerado pelo procedimento.

```
FOOTNOTE "<rodapé>" ;
```

[63]