Data Entry Using SAS [®]- A Discussion of Various Possibilities

Ass. Prof. Dipl.Ing. Dr. Barbara Schneider
Institute for Medical Statistics, University of Vienna,Vienna, Austria

ABSTRACT

This paper is somehow a summary of experiences I have made with data entry when introducing SAS to students in medicine or to physicians. Many of them preparing their theses or doing scientific work are involved in data collection and data analysis. An important phase between data collection and data analysis is data entry.

Most books concerning data analysis using SAS demonstrate data entry only in combination with the data step using the cards or datalines statement. Nowadays hardly someone can be convinced in typing in data in that ancient way.

This paper introduces more attractive possibilities of data entry and should serve as a guide for SAS novices. Advantages and disadvantages of data entry using VIEWTABLE, SAS Analyst, SAS/INSIGHT and the FSEDIT Procedure (SAS/FSP) are discussed. Recommendations are provided to help choosing the most efficient strategy for data entry and suggestions to improve some features are made. Topics as data creation, adding rows and columns to data sets or computing columns are covered as well. Above all the import facility for Microsoft[®] EXCEL and ACCESS data sets will be highlighted.

SAS Software refers to Version 8.2 for Microsoft[®] Windows

INTRODUCTION

During planning phase of a medical study case report forms (CRF) are designed. Roughly speaking all information of interest (variables) are documented in a structured way. During ongoing of the study patients characteristics (variables) are measured and these informations are stored in the specific identified copy of the CRF for each patient. Simultaneously data tables have to be created to store the informations electronically. Before creating a data set one must be aware of the nature of the variables. Variables can be categorial (nominal or ordinal) or metric (discrete or continuous). It is recommendable that you list your variables in the order of entering data - the order of variables in the CRFs.

To create a data table you first have to know which variables (columns) to incorporate and then the type of the specific variables. Using SAS Software you have to distinguish between type numeric and character. Only for numeric variables you can perform arithmetic operations. Type character is suitable for categorial variables especially for nominal variables. But in some cases it is advantageous to use type numeric even for nominally scaled variables. Date variables such as birthdate are of type numeric. Formats and Informats manage to input and display date variables in the desired way.

Throughout the paper a simple data set for demonstrations is used. This data table meets the needs to show, compare and discuss the various possibilities of data creation and data entry using SAS.

Sample data set

Variable list:

|	
PNR	Patient's identification number
BDATE	Date of birth
EDATE	Date of examination
SEX	Gender (f=female , m=male)
RRS	Systolic blood pressure
RRD	Diastolic blood pressure
PAIN	Pain (0=no pain , 1= bearable pain, 2= unbearable pain)

The first five patients are recorded as:

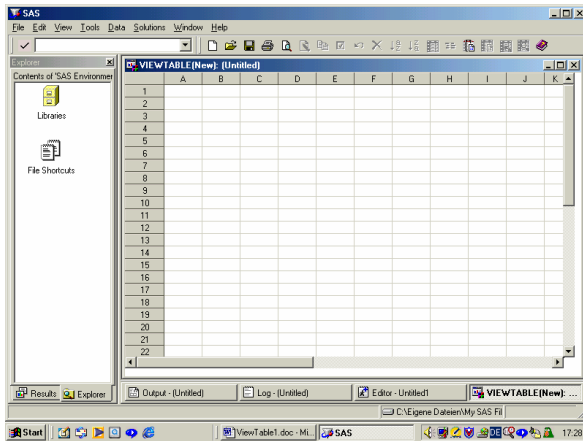| PNR	BDATE	EDATE	SEX	RRS	RRD	PAIN
1	23/03/1950	05/12/2000	m	140	85	1
2	04/10/1972	11/12/2000	m	120	70	0
3	17/08/1943	14/12/2000	f	200	90	0
4	09/07/1977	08/01/2001	m	135	80	2
5	21/02/1963	09/01/2001	f	115	70	1

Now let's create this data table using

VIEWTABLE

To create a new table select **Tools** from the menu and click on **Table Editor**
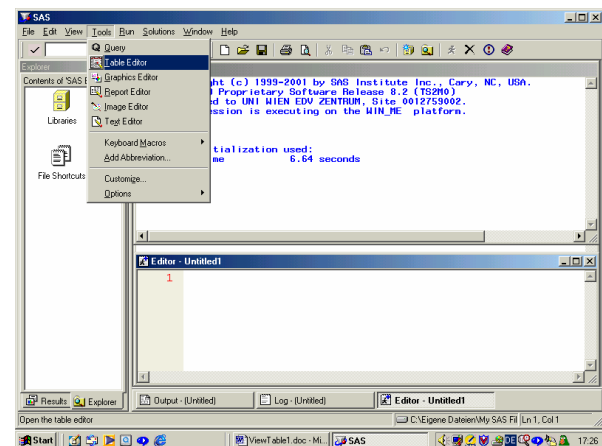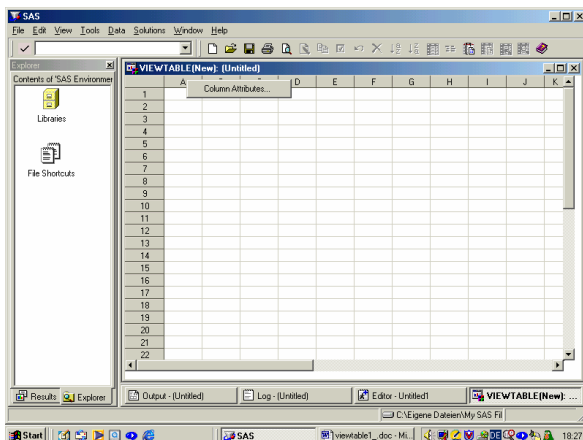

An empty VIEWTABLE window opens

To define the structure of a data table that means defining columns, you have to set attributes of each column, such as its label, whether it is a character or numeric column, and its Format or Informat.
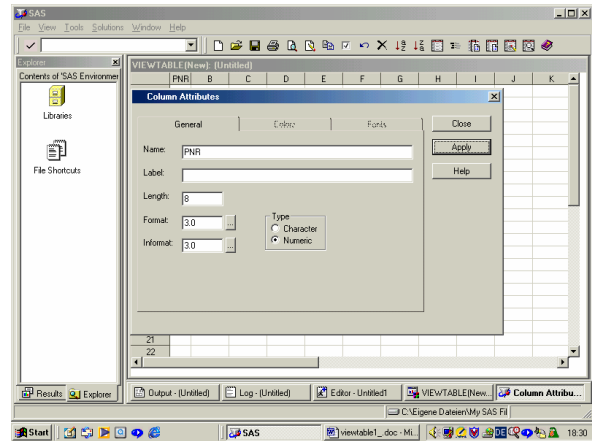
To set the first column right-click on **A** then select **Column Attributes**

Drag and drop the Column Attributes Dialog Box so that you can view the column headings A, B, C,.. then fill in the fields

Name: PNR
Label: Patientidentification
Length: 8 (this value is preset)
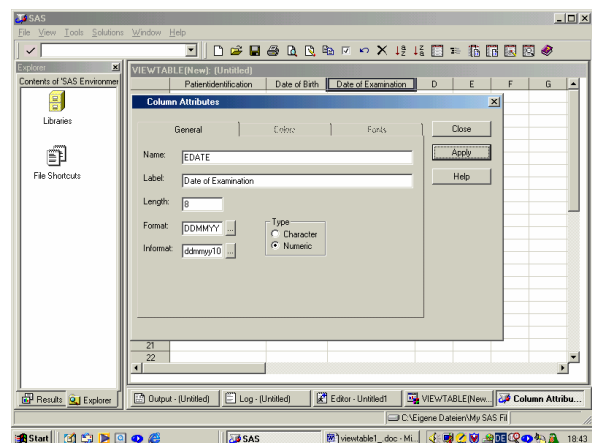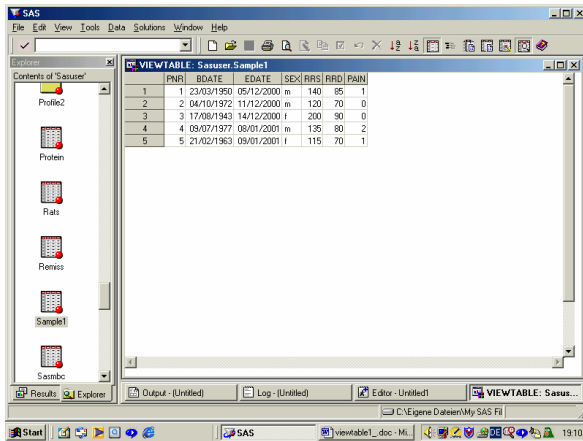Format: 3.0
Informat: 3.0

Type is set to Numeric

After completing the fields by entering Patientidentification as a label click the Apply button. Select column **B** and proceed in the same way as before
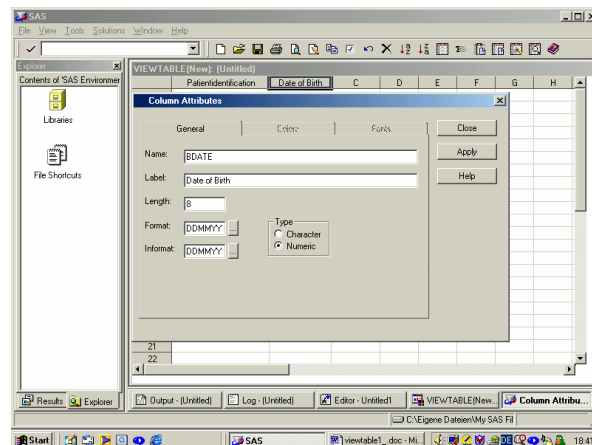

You must not forget to click **Apply** to set the attributes. The following tables display the various column definitions.

Finally close the Column Attributes Dialog Box. Labels are shown as column headings. To have column names displayed instead of labels select **Column Names** from the **View** pop-up menu

After completing the first five records save the data table selecting **File** → **Save As** and then close the Viewtable window. To open the existing table you can use the SAS Explorer window: Expand the library containing the table you want to open. Double-click the table name

Now you can add more records to the data set. If you open a data table you are in brows mode. To add rows you first have to change to edit mode by selecting **Edit** from the **Edit** pop-up menu
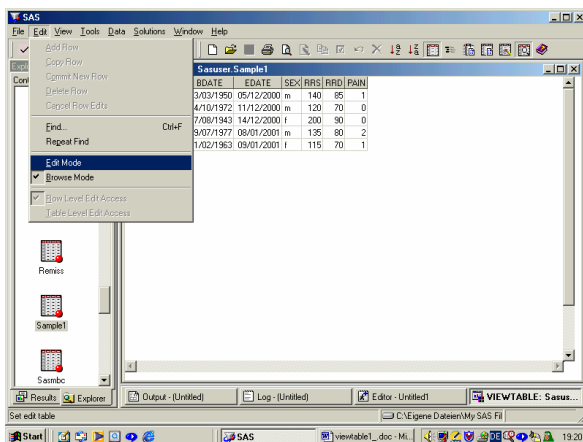
Then select **Add Row** from the **Edit** pop-up menu
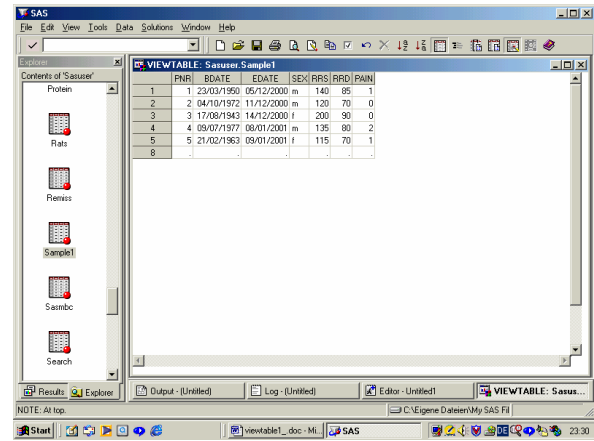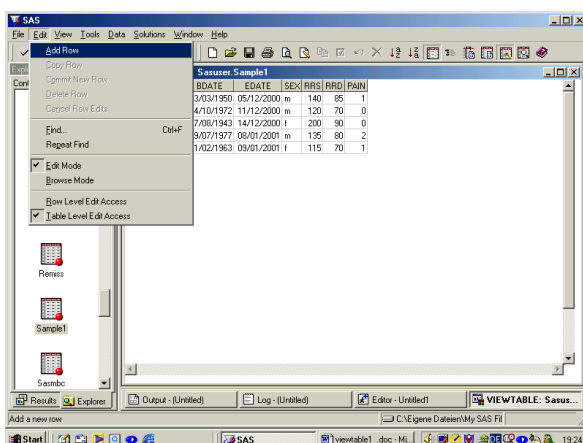
Using Viewtable you can easily create a data table by setting the columns, you can add rows to the table but you can't add columns nor you can't compute columns from existing columns. This is a considerable disadvantage .

DATA CREATION AND DATA ENTRY USING THE ANALYST APPLICATION

The Analyst application is a data analysis tool within SAS Software designed not only to provide easy access to basic statistical analyses but also to data creation and data entry. The application is intended for students and novice SAS users as well as for researchers and analysts who are already familiar with the SAS System.

To start the Analyst application simply select **Analysis** from the **Solutions** pop-up menu and click on **Analyst**
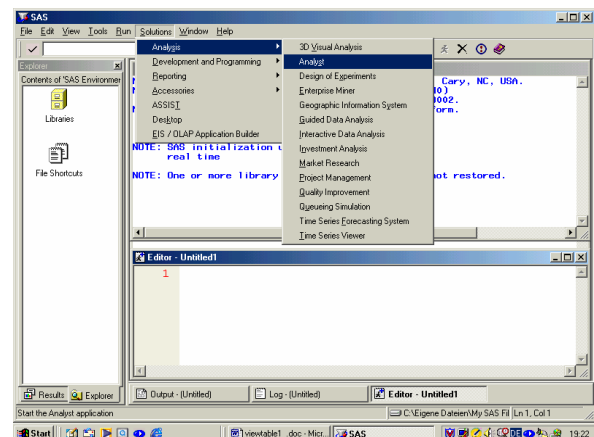
If there is no project stored an emty data table is displayed else select **File → New** to create a new data table.
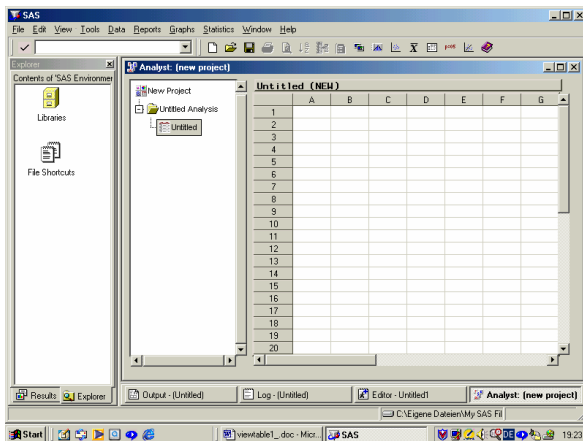
Compared to the Viewtable window you can't define column attributes in New mode. This is a disadvantage because it isn't possible setting Informats and Formats and this causes problems especially with date variables. Click on the column headings to assign column names. Using the Analyst you can add columns at any position you like and you can compute columns. This is an advantage compared to Viewtable.
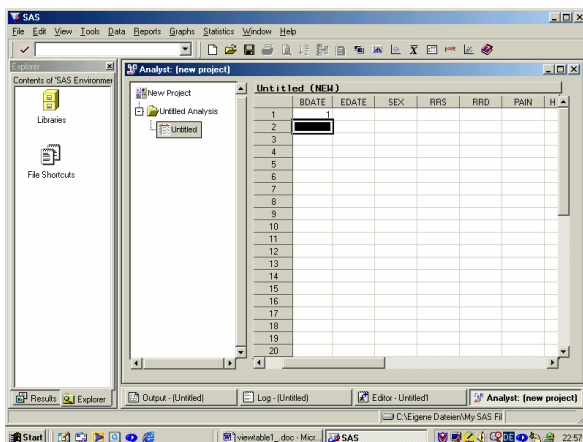
Entering a number into the first field below the column name will assign type numeric to the entire column otherwise the column is set to type character. That means if you enter a date value e.g. 23/03/1950 this value is interpreted as type character. There are several possibilities to overcome this problem. If the table is stored and you open the table again, change to edit mode then it is possible to assign column attributes by right-clicking the mouse button. Notice within Analyst column attributes are called Properties. So one strategy to define column properties can be: create one column with one observation, save the table and reopen the table. Then you are able to add columns in edit mode and set type, format and informat of the column.
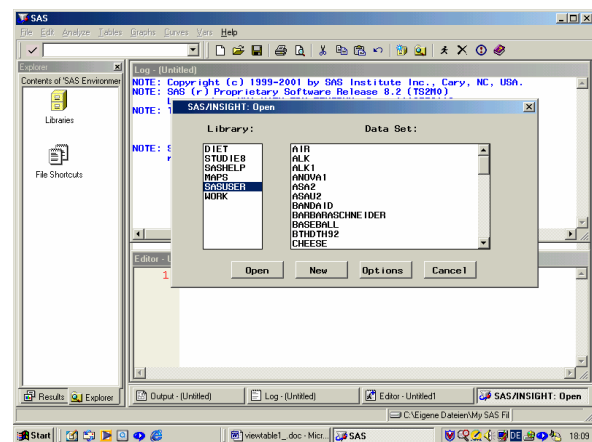
DATA CREATION AND DATA ENTRY USING SAS/INSIGHT

Primarily SAS/INSIGHT software is a tool for data exploration and analysis but it is aswell nice for data creation and data ntry.
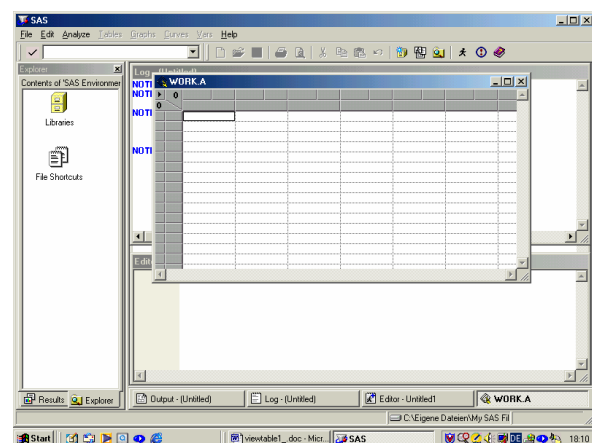
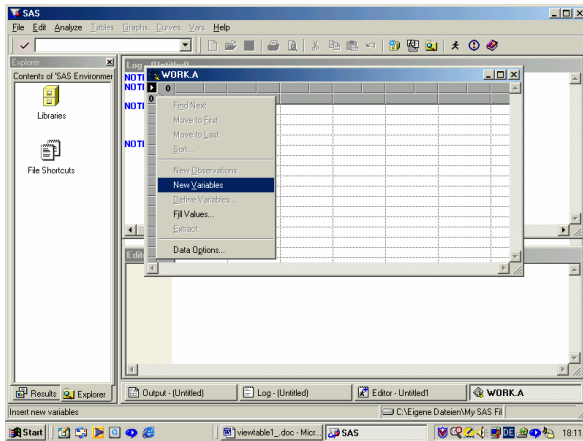To invoke SAS/INSIGHT select **Solutions → Analysis → Interactive Data Analysis**
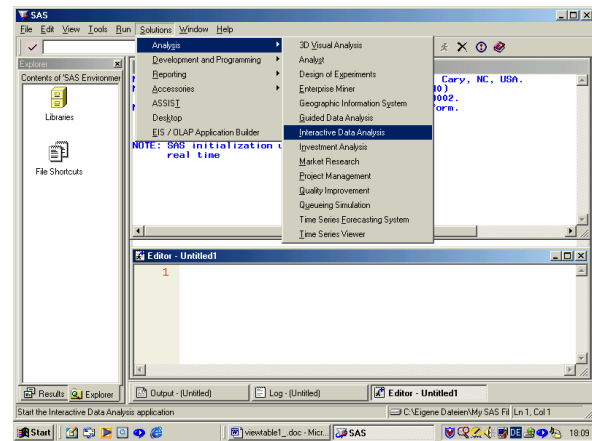
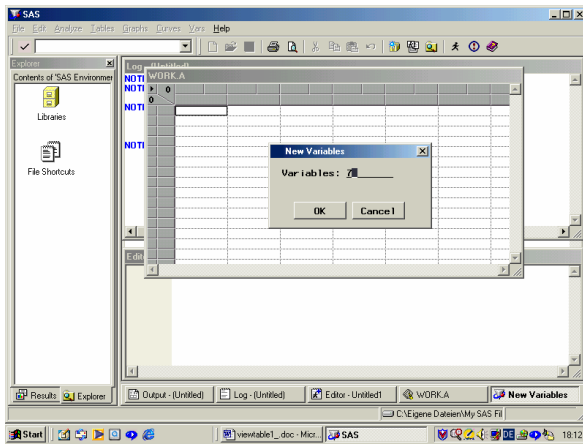Click on **New** to create a new data table

An empty data table is displayed

To create variables and assign types and measurement levels click in the upper left corner of the data window. This displays the data pop-up menu. Choose **New Variables** from the pop-up menu.
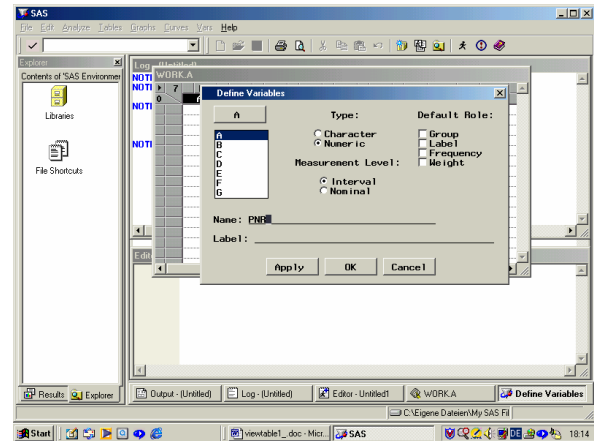
This displays a dialog to prompt you for the number of variables to create.


Enter "7" in the **New Variables** dialog, then click **OK**. The data window should appear as shown in the next figure.
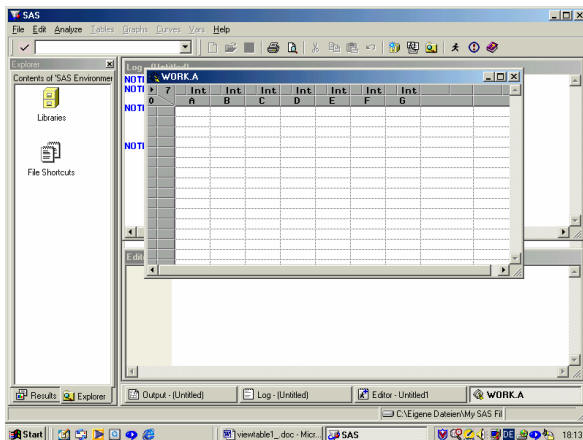
Double-click on **A** to start defining columns

To assign Formats (next table) select **Edit** → **Formats** → **Other** (if necessary e.g. for date Formats). Informats are set to Formats (Caution: there seems to be a bug using date formats in version 8.2). Compared to Viewtable and the Analyst application you can't move or hide columns but fortunately you can add columns.
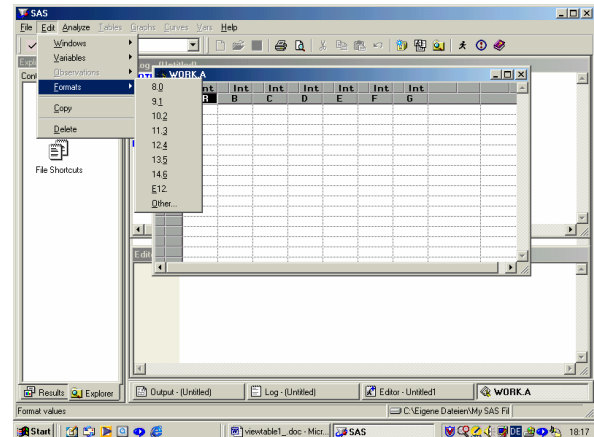
USING THE FSEDIT PROCEDURE FOR DATA CREATION AND DATA ENTRY

The FSEDIT procedure enables you to create and edit SAS data sets in an interactive windowing environment.

To create a new data table submit the following code:
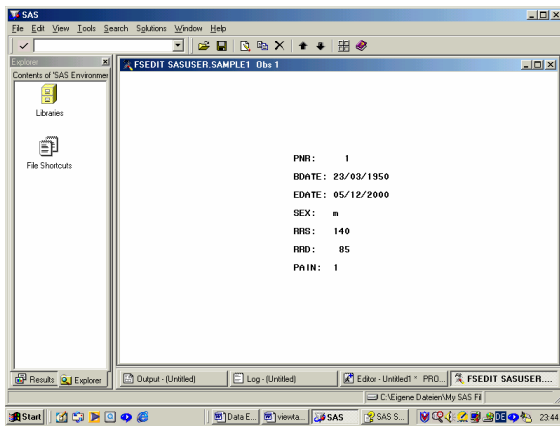
proc fsedit new = *SAS-data-set***;**
run;

When you invoke the FSEDIT procedure using the NEW = option in the PROC FSEDIT statement, you enter data set definition mode.
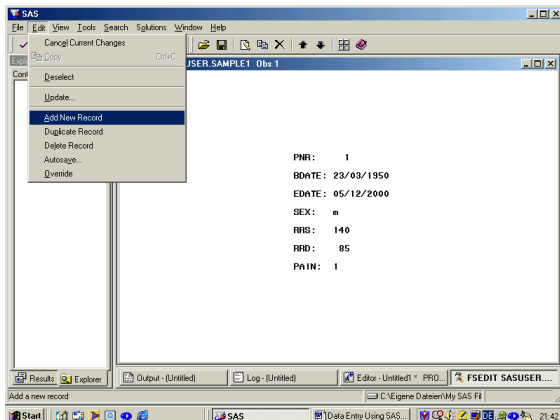
Submitting the above SAS code will display the following window.

6
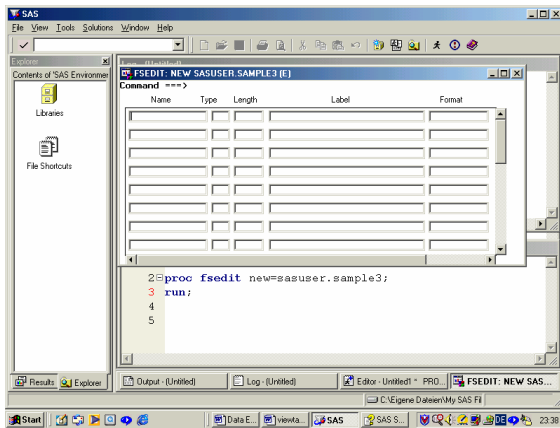
After assigning type, length, label, formats and informats use the end command to open the new data table windoe then enter the first record

To add a new record select **Add New Record** from the **Edit** pop-up menu

A disadvantage using FSEDIT is that you can't add new colums to the table.

FINAL REMARKS

If you are familiar in writing SAS code the simplest and most efficient way creating a new data table is by writing a SAS program like this:

```
data sasuser.sample1;

format
PNR    3.
BDATE  ddmmyy10.
EDATE  ddmmyy10.
SEX    $1.
RRS    3.
RRD    3.
PAIN   1.
;
informat
PNR    3.
BDATE  ddmmyy10.
EDATE  ddmmyy10.
SEX    $1.
RRS    3.
RRD    3.
PAIN   1.
;
run;
```

The advantage of proceeding in that way is that you can add columns at any position even if there are already observations entered. For example add the numeric variable GOT before RRS

Caution: don't forget the **set** statement!

```
data sasuser.sample1;
format
PNR    3.
BDATE  ddmmyy10.
EDATE  ddmmyy10.
SEX    $1.
GOT    5.
RRS    3.
RRD    3.
PAIN   1.
;
informat
PNR    3.
BDATE  ddmmyy10.
EDATE  ddmmyy10.
SEX    $1.
GOT    5.
RRS    3.
RRD    3.
PAIN   1.
;
set sasuser.sample1;
run;
```

I don't want to finish without mentioning **Enterprise Guide.** Apart from many nice features Enterprise Guide is aswell a very recommendable tool for data creation and data entry. As far as data creation is concerned Enterprise Guide combines both the advantages of Viewtable and the Analyst application: e.g. defining type, formats and informats in new mode, add columns, compute columns, import and export data tables.

To import data from Microsoft® Excel or Access you can use both Viewtable and SAS Analyst.

Using Viewtable select **File → Import Data**, the Import Wizard opens.

7
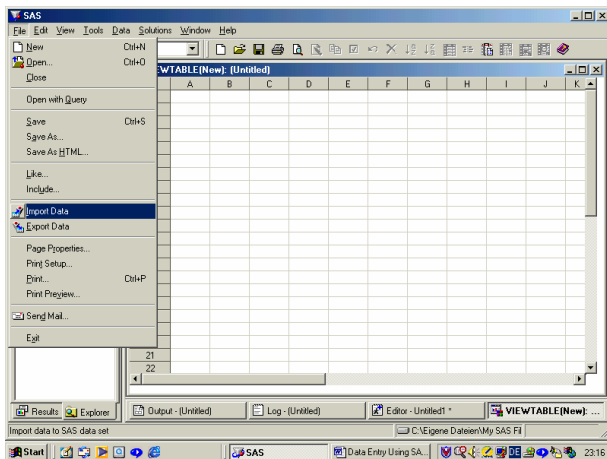
SAS, SAS Software, Base SAS, SAS/INSIGHT, SAS/FSP, the Analyst Application software and the Enterprise Guide are registered trademarks of SAS Institute Inc., Cary, NC, USA

REFERENCES

SAS ® Language: Reference, Version 8
Getting Started with the SAS® System, Version 8
The Analyst Application®, First Edition
SAS/INSIGHT® User's Guide, Version 8,
SAS/FSP® Software Procedures Guide, Version 8
Getting Started with Enterprise Guide® Software, First Edition

ADDRESS FOR CORRESPONDENCE

Ass. Prof. Dipl.Ing. Dr. Barbara Schneider
Institute for Medical Statistics
University of Vienna
Schwarzspanierstr. 17
A - 1090 Vienna
Austria
barbara.schneider@univie.ac.at

Importing data using Analyst select **File → Open →** select the desired data type then click Open.