

Paper 183-2009

A Discussion of Various Possibilities for Data Entry Using SAS® - An Evergreen

Barbara G. Schneider, Medical University of Vienna, Austria
 Kristan A. Schneider, University of Vienna, Austria

ABSTRACT

This video poster is a summary of experiences we have made with data entry when introducing SAS to medical students or physicians. For preparing their theses or doing scientific work they are involved in data collection and data analysis. An important phase between data collection and data analysis is data entry, the topic to which this paper is devoted to. Most books concerning data analysis using SAS demonstrate data entry only in combination with the data step using the cards or datalines statement. Nowadays, however, hardly someone can be convinced to enter data in this 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, Enterprise Guide, SAS/INSIGHT, SAS Stat Studio 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 such 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 9.2 for Microsoft® Windows

INTRODUCTION

During planning phase of a medical study case report forms (CRF) are designed. Roughly speaking, all information of interest (variables) is documented in a structured way. During ongoing of the study patients characteristics (variables) are measured and these information are stored in the specific identified copy of the CRF for each patient. Simultaneously data tables have to be created to store the information electronically. Before creating a data set one must be aware of the nature of the variables. Variables can be categorical (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 need to know which variables (columns) to incorporate and of which type they are. Using SAS Software you have to distinguish between two types, namely, *numeric* and *character*. Only numeric variables are designed to perform arithmetic operations. The numeric type subsumes date variables, e.g., date of birth, which are managed to be entered and displayed in the desired way by choosing appropriate *INFORMAT* and *FORMAT* options, respectively. The character type is designed for categorical and especially for nominal variables. However, in some it is preferable to use the numeric type for nominally or categorically scaled variables.

Throughout the paper we use the same simple data set, which is sufficient to demonstrate, compare and discuss the various possibilities of data creation and data entry using SAS.

Sample data set

List of variables:

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)

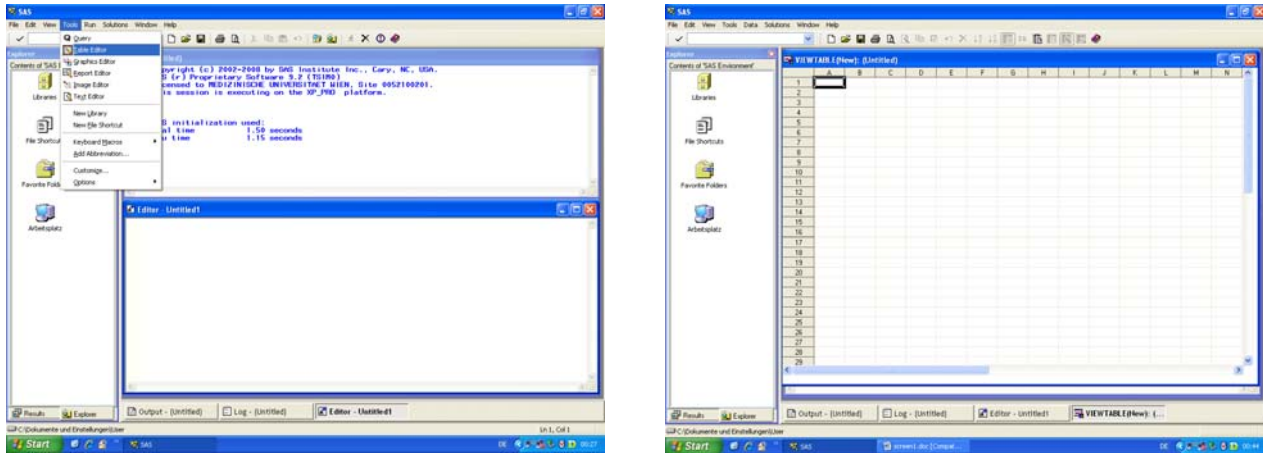
The first five patients are recorded as:

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

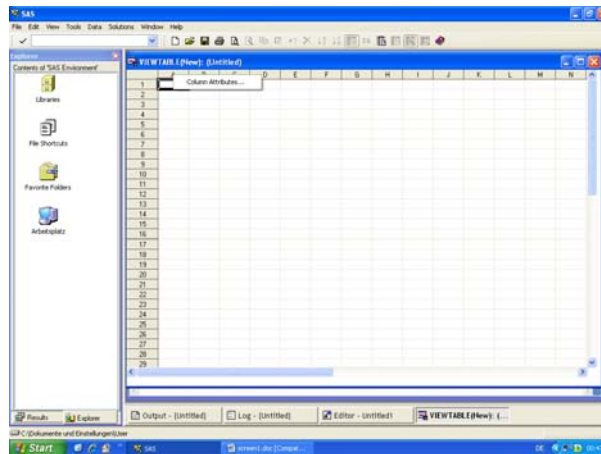
First, we create this data table using *Viewtable*.

VIEWTABLE

To create a new table select **T**ools from the menu and click on **T**able Editor. An empty VIEWTABLE window opens.



To define the structure of a data table, i.e., 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** and then select **Column Attributes**.

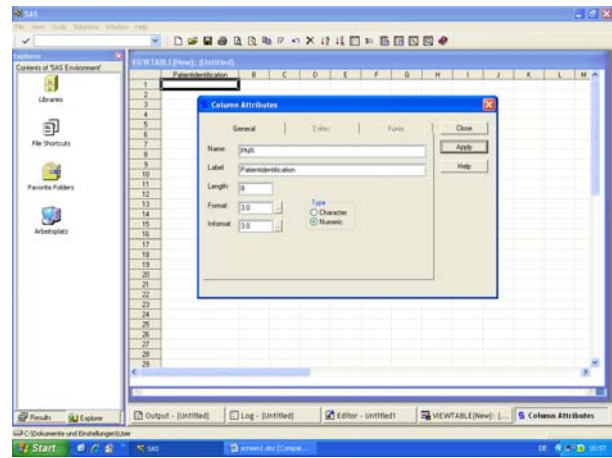
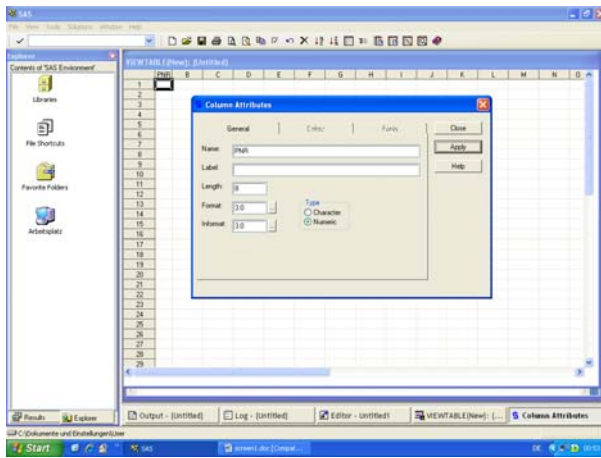


Drag and drop the column attributes dialog box so that you can view the column headings A, B, C, etc. Next, fill in the fields

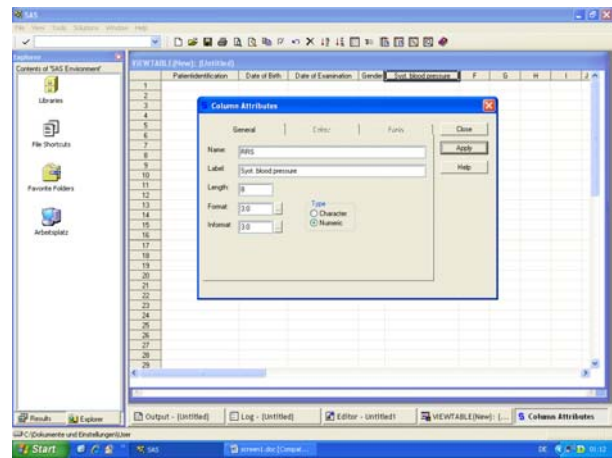
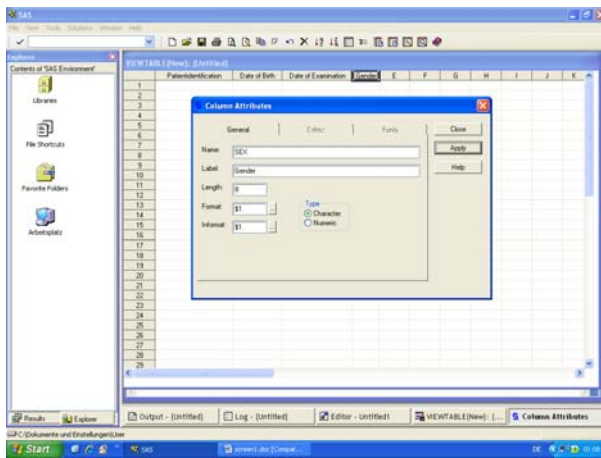
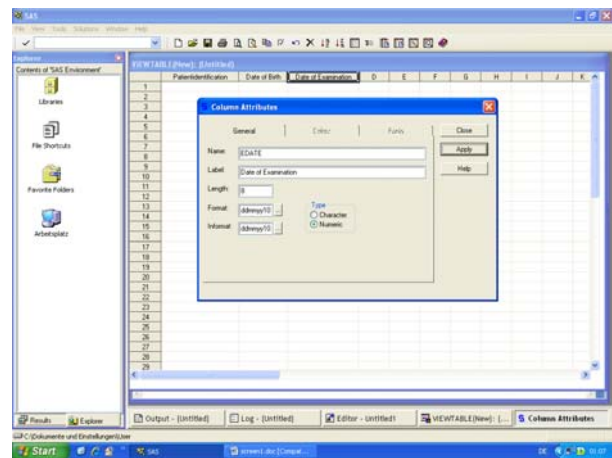
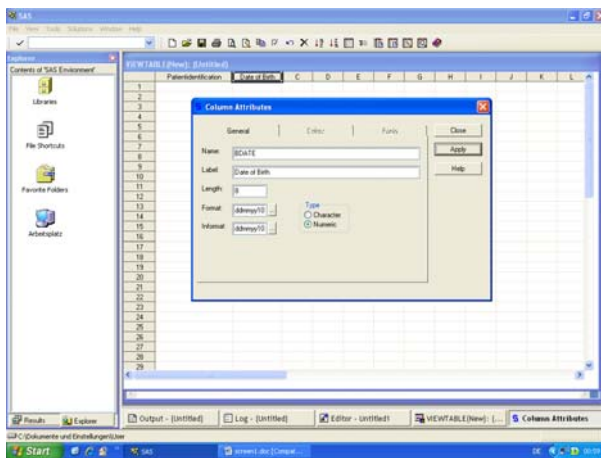
Name: PNR
 Label: Patientidentification
 Length: 8 (this value is preset)
 Format: 3.0
 Informat: 3.0

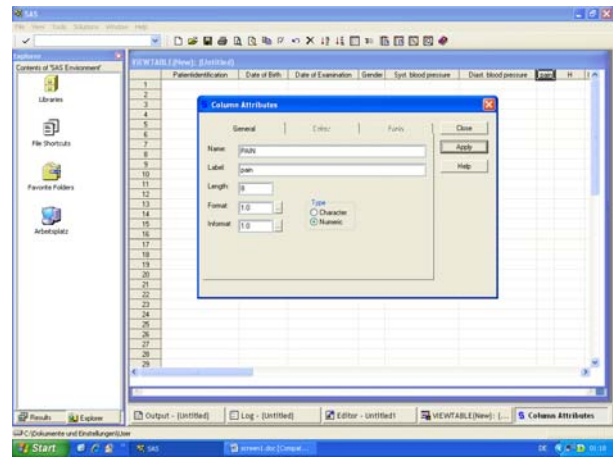
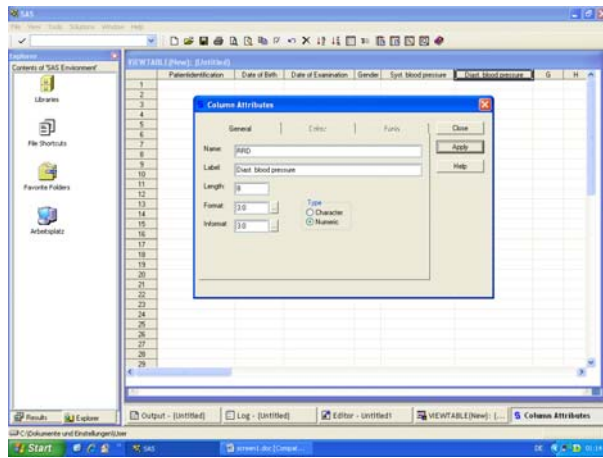
Type is set to **Numeric**.

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

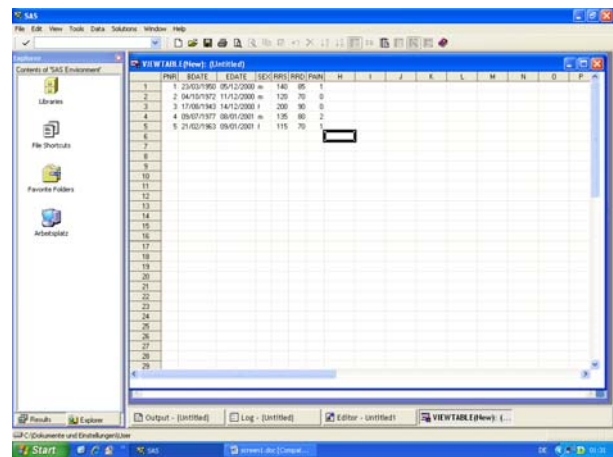
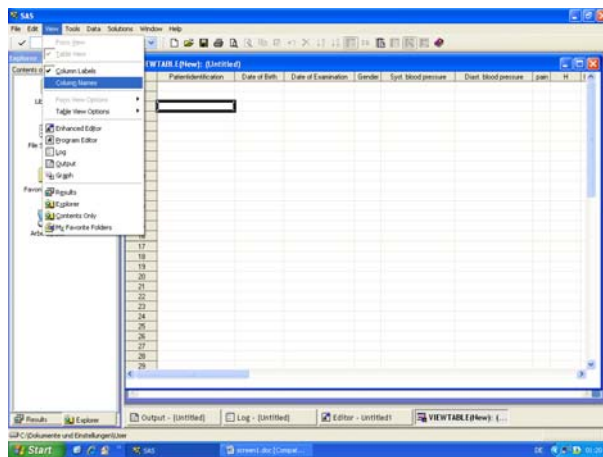


You must not forget to click **Apply** to set the attributes. The following tables display the various column definitions.

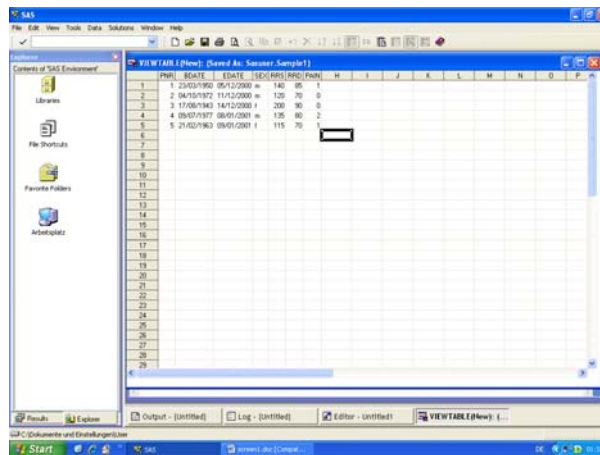




Finally, close the **Column Attributes** dialog box. By default labels are shown as column headings. To display column names instead of labels you have to select **Column Names** from the **View** pop-up menu.

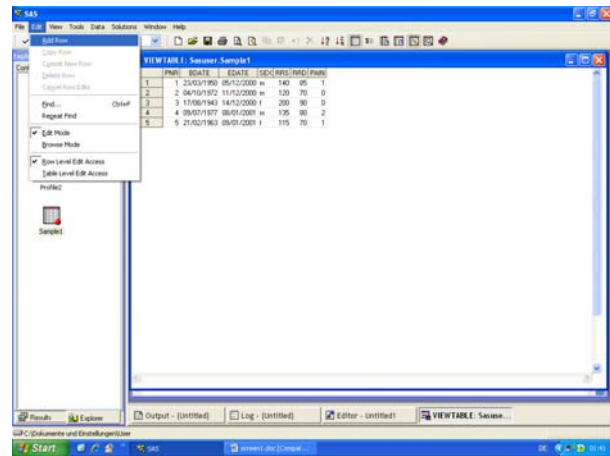
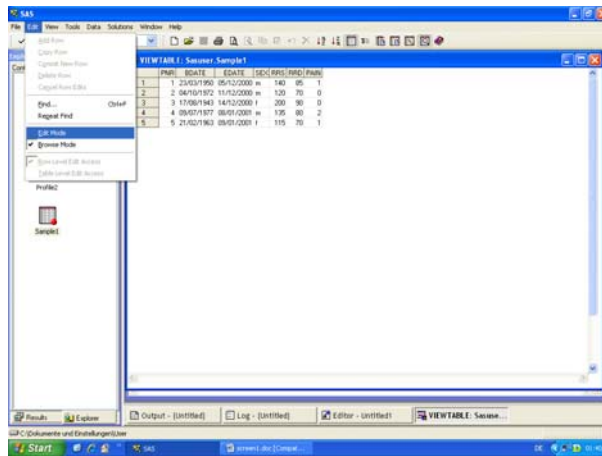


After completing the first five records save the data table selecting **File** → **Save As** and then close the Viewtable window.

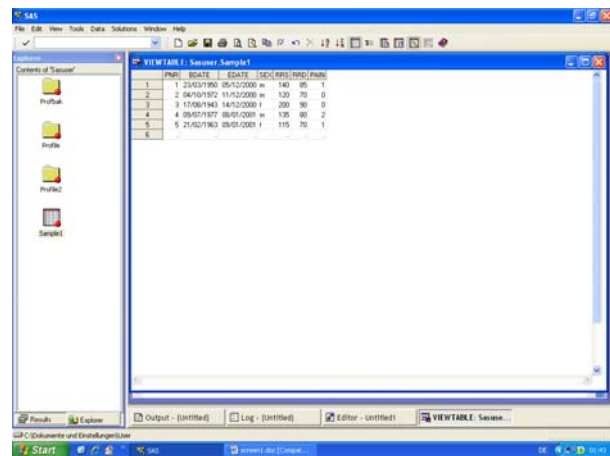
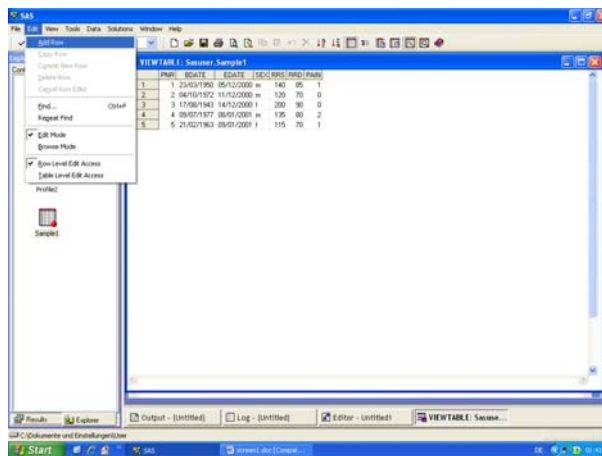


To open the existing table you can use the SAS Explorer window: expand the library containing the table you want to open and double-click the table name.

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 switch to edit mode by selecting **Edit** from the **Edit** pop-up menu.



Afterwards, select **Add Row** from the **Edit** pop-up menu.

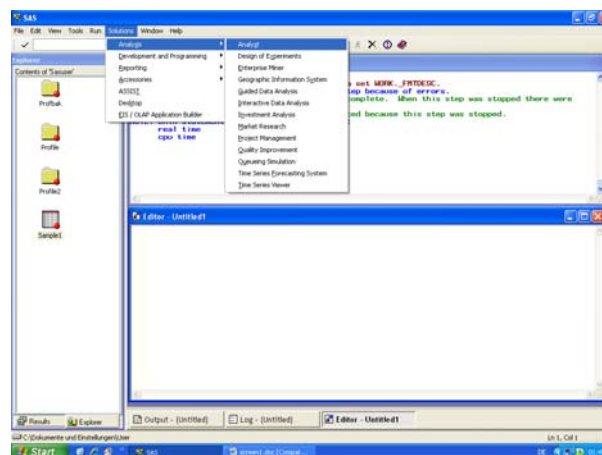


Using Viewtable you can easily create a data table by setting the columns, you can add rows to the table. However, neither can you add columns, nor can you compute columns from existing ones. This is a considerable disadvantage.

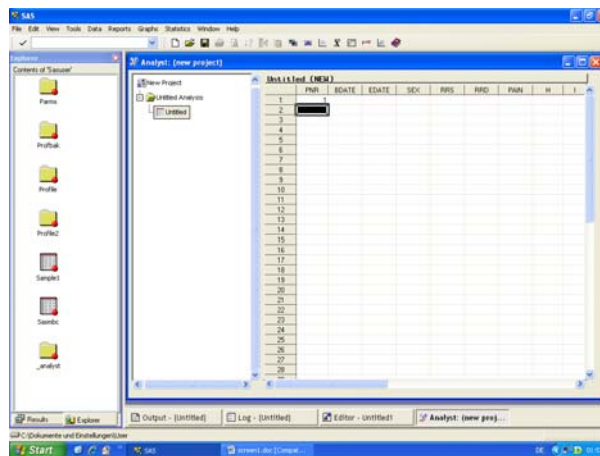
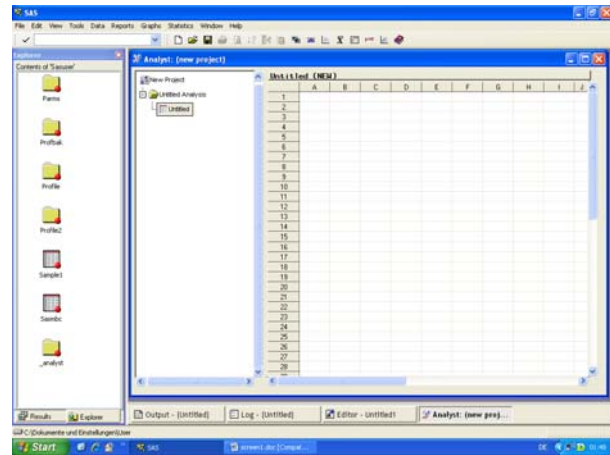
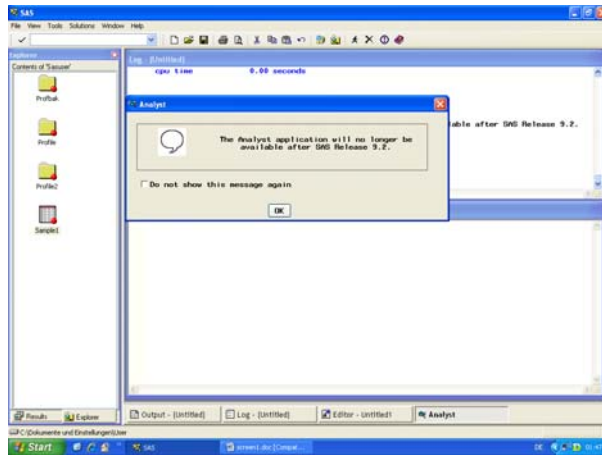
DATA CREATION AND DATA ENTRY USING THE ANALYST APPLICATION

The *Analyst application* is a point-and-click based 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 select **Analysis** from the **Solutions** pop-up menu and click on **Analyst**. If there is no project stored, an empty data table should be displayed. Otherwise select **File** → **New** to create a new data table.



Compared to the Viewtable window you cannot define column attributes in **New** mode. This is a disadvantage because it is impossible to set INFORMATs and FORMATS. Especially with date variables this causes problems. 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 from existing ones, which is an advantage compared to Viewtable.



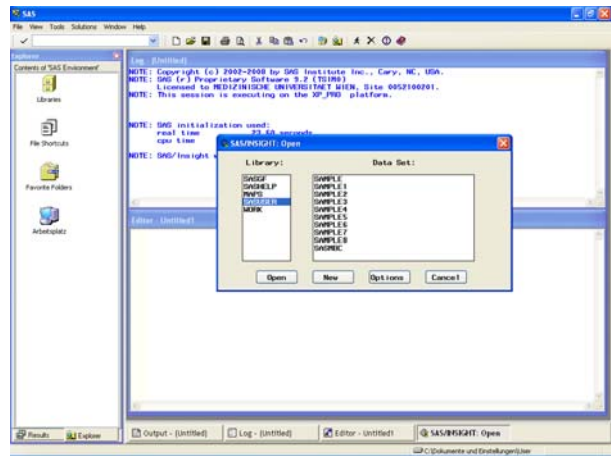
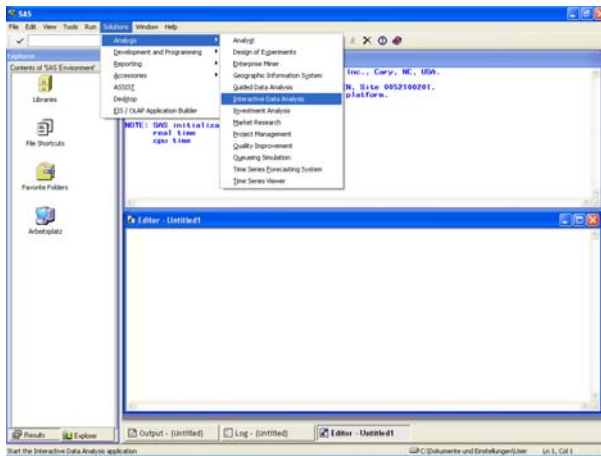
Entering a number into the first field below the column name will assign type numeric to the entire column otherwise the column is set to type character. That means if you enter a date value, e.g. 23/03/1950 this value is interpreted as a character. There are several possibilities to overcome this problem. If the table is stored and you open the table again, it is possible to assign column attributes by right-clicking the mouse button after changing to edit mode. Notice that within Analyst column attributes are called **Properties**. Hence, in order to define column properties, one could proceed as follows. Create one column with one observation and save this table. After reopening the table you are able to add columns in edit mode and set TYPE, FORMAT and INFORMAT of the column.

Although the **Analyst Application** is still available in SAS Version 9.2, it will no longer be available in subsequent versions. It will be replaced by the meanwhile well approved and accepted Enterprise Guide. A new version, Enterprise Guide 4.2, is scheduled for 2009.

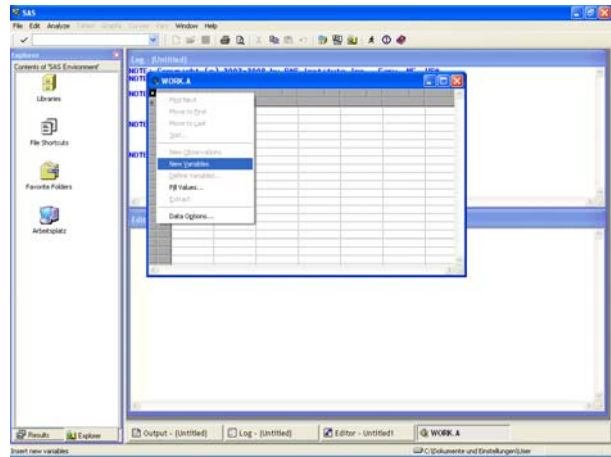
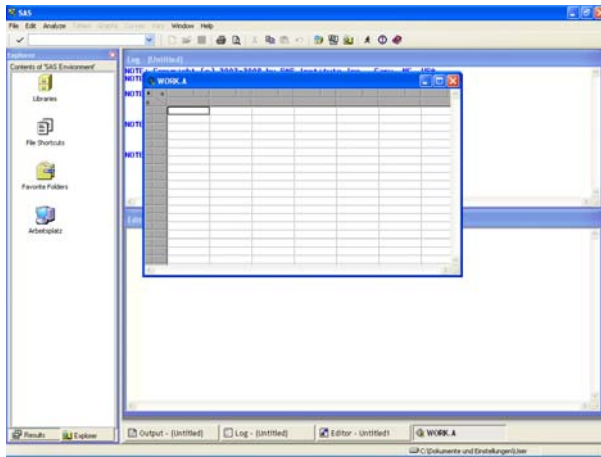
DATA CREATION AND DATA ENTRY USING SAS/INSIGHT

Primarily SAS/INSIGHT software is a tool for data exploration and analysis but it is also a useful tool for data creation and data entry.

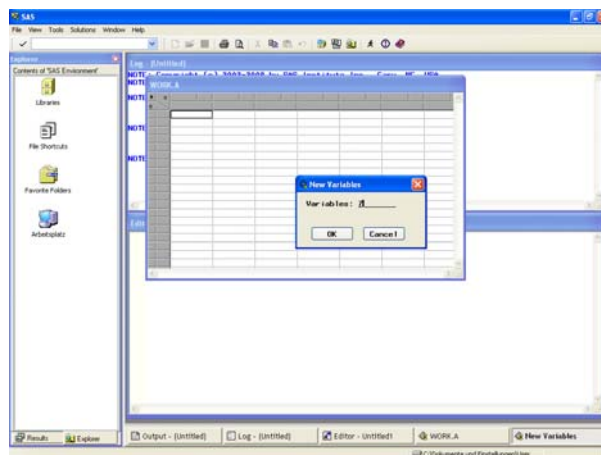
To invoke SAS/INSIGHT select **Solutions** → **Analysis** → **Interactive Data Analysis**. Click on **New** to create a new data table.



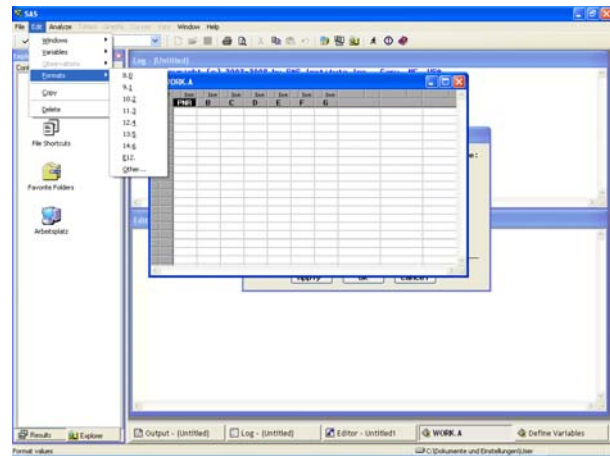
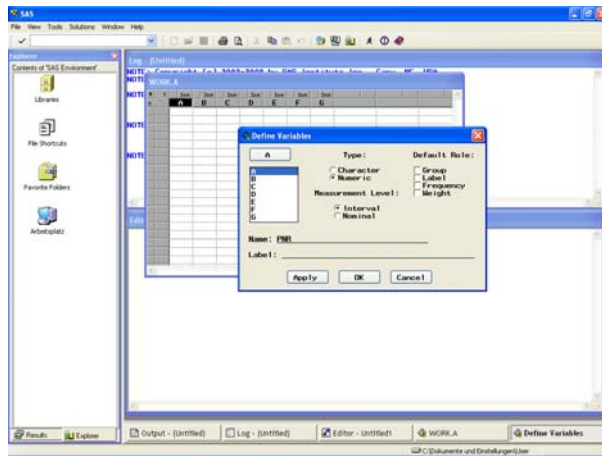
An empty data table is displayed. To create variables and assign types and measurement levels click in the upper left corner of the data window. This displays the data pop-up menu. Choose **New Variables** from the pop-up menu.



This displays a dialog to prompt you for the number of variables to create. Enter "7" in the **New Variables** dialog and 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 if someone still uses this version). Compared to Viewtable and the Analyst application you cannot move or hide columns but fortunately you can add columns.



SAS/INSIGHT is replaced by SAS Stat Studio for future Versions.

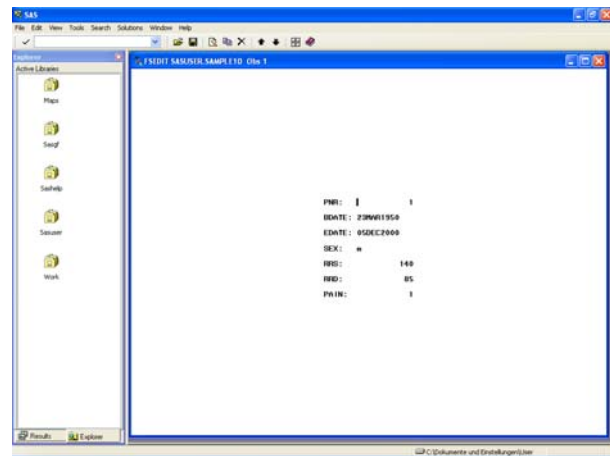
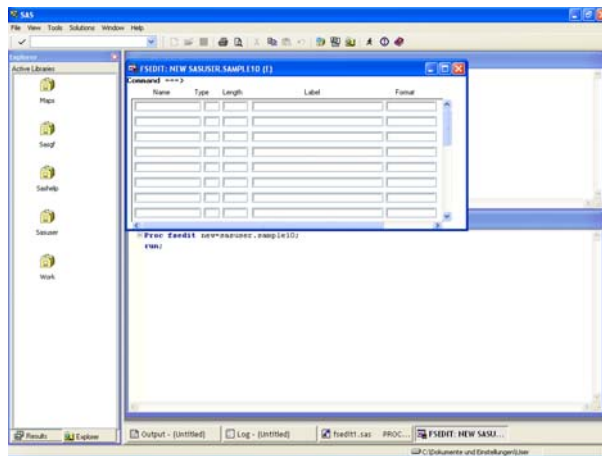
USING THE FSEDIT PROCEDURE FOR DATA CREATION AND DATA ENTRY

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

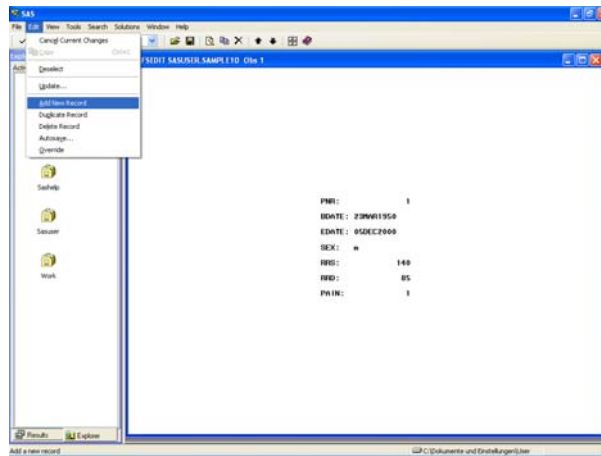
To create a new data table submit the following code:

```
proc fseedit new = SAS-data-set;  
run;
```

When you invoke the FSEDIT procedure using the NEW = option in the PROC FSEDIT statement, you enter the data-set definition mode. Submitting the above SAS code will display the following window. After assigning TYPE, LENGTH, LABEL, FORMAT and INFORMAT use the end command to open the new data table window. 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 cannot add new columns 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 not 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;

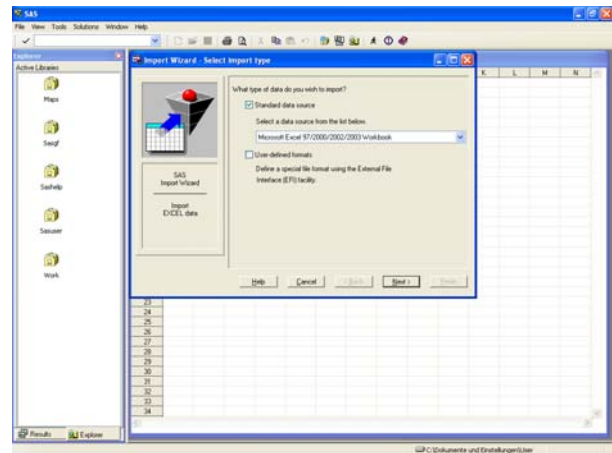
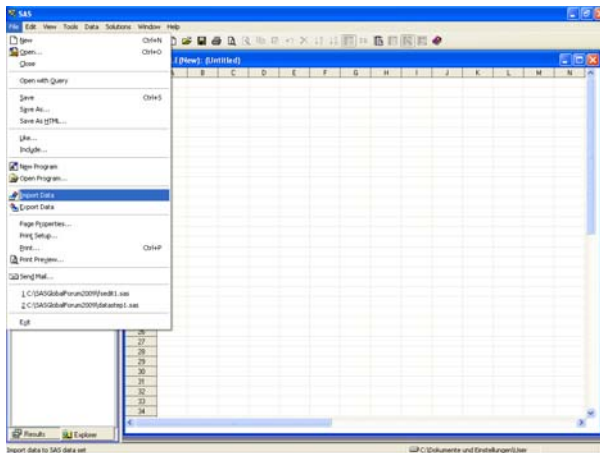
```

We do not want to finish without mentioning the benefits of **Enterprise Guide**. Apart from many useful features Enterprise Guide is also 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 informat in new mode, add columns, compute columns, import and export data tables.

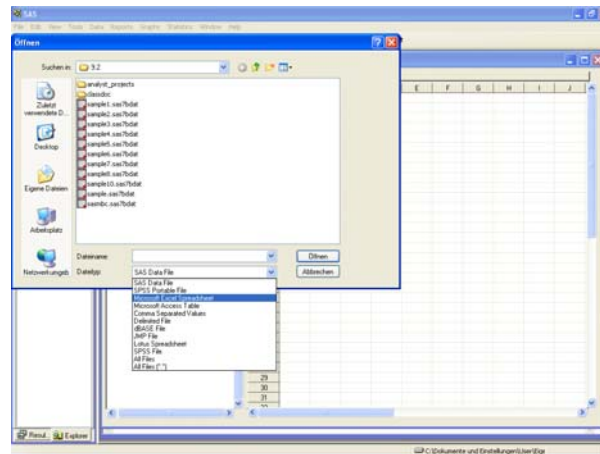
IMPORTING DATA FROM AN EXTERNAL SOURCE

To import data from Microsoft[®] Excel or Access you can use both Viewtable, SAS Analyst Enterprise Guide and SAS Stat Studio.

Using Viewtable select **File** → **Import Data**. The Import Wizard opens.



To import data using Analyst select **File** → **Open** and select the desired data type. Afterwards click **Open**.



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 9.2
 Getting Started with the SAS[®] System, Version 9.2
 The Analyst Application[®], First Edition
 SAS/INSIGHT[®] User's Guide, Version 9.2,

SAS/FSP® Software Procedures Guide, Version 9.2
Getting Started with Enterprise Guide® Software,
SAS Stat Studio 3.1: User's Guide
Data Entry Using SAS ®- A Discussion of Various Possibilities, SUGI 27 paper

ADDRESS FOR CORRESPONDENCE

Ass. Prof. Dipl.Ing. Dr. Barbara Schneider
Institute for Medical Statistics
University of Vienna
Spitalgasse 23
A - 1090 Vienna
Austria
barbara.schneider@meduniwien.ac.at