PhUSE 2015

Paper IS08

Don't be afraid of PROC REPORT – a step-by-step guide

Nicola Tambascia, Accovion GmbH, Eschborn, Germany
Konstanze Morgenroth, Accovion GmbH, Eschborn, Germany

ABSTRACT

As a statistical programmer in the pharmaceutical industry you will come across PROC REPORT sooner or later during your daily work. When first looking at the structure of this powerful procedure you might be overwhelmed by the various options it offers. This paper will get you started with PROC REPORT and will help you to understand how to use the different options available to easily create the outputs you need.

This paper is intended to provide a brief introduction to the SAS® PROC REPORT procedure for beginners. SAS® Version 9.3 was used as basis for the content of this paper.

SIMPLE PROC REPORT

PROC REPORT can be used to create a simple report without having a deeper knowledge of the procedure. The basic structure of PROC REPORT is the following (a title was added by using the title statement before the PROC REPORT):

```
title 'Simple report';
proc report data = adsl;
run;
```

All variables of the selected dataset (in this example a minimized demography dataset was used) are listed. Submitting this simple statement, the output looks like this:

Simple report 17:09 Wednesday, September 9, 2015 1

Date of Sign. on Informed Consent (SAS date)	Sex (DVG Number)	Race (DVG Number)	Ethnicity (DVG Number)	Age	STUDYID	AGEGRP	SEX	RACE	ETHNIC	VISIT	USUBJID
16NOV2012	1	1	2	17	Test Study	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	V1: ENROLLMENT	1010001
15NOV2012	1	1	2	11	Test Study	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	V1: ENROLLMENT	1010002
16NOV2012	1	1	2	15	Test Study	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	V1: ENROLLMENT	1010003
16NOV2012	2	1	2	9	Test Study	Group 2 (6 - < 12 years)	Female	White	Not Hispanic or Latino	V1: ENROLLMENT	1010004
16NOV2012	2	1	2	16	Test Study	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	V1: ENROLLMENT	1010005
19NOV2012	1	1	2	17	Test Study	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	V1: ENROLLMENT	1010006
19NOV2012	2	1	2	16	Test Study	Group 1 (12 - < 18 years)	Female	Asian	Not Hispanic or Latino	V1: ENROLLMENT	1010007
19NOV2012	1	1	2	16	Test Study	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	V1: ENROLLMENT	1010008
19NOV2012	1	1	2	14	Test Study	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	V1: ENROLLMENT	1010009
19NOV2012	2	1	2	17	Test Study	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	V1: ENROLLMENT	1010010

SELECTING VARIABLES

As probably not all variables are required in the output, the next step should be to choose the variables to be displayed. This is easily done by adding the COLUMNS statement to the code:

```
title 'Selecting variables';
proc report data = adsl;
   columns studyid usubjid agegrp sex race ethnic age;
run;
```

The output below shows that only the variables specified in the COLUMNS statement are displayed. Please note: the order in which you state the variables in the COLUMNS statement will determine the order of the variables in the output:

Selecting variables 17:09 Wednesday, September 9, 2015 1

STUDYID	USUBJID	AGEGRP	SEX	RACE	ETHNIC	Age
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	11
Test Study	1010003	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	15
Test Study	1010004	Group 2 (6 - < 12 years)	Female	White	Not Hispanic or Latino	9
Test Study	1010005	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	15
Test Study	1010006	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17
Test Study	1010007	Group 1 (12 - < 18 years)	Female	Asian	Not Hispanic or Latino	16
Test Study	1010008	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	16
Test Study	1010009	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	14
Test Study	1010010	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	17

DEFINING VARIABLES

Now that the required variables are selected, the appearance of the report can be enhanced. This can be achieved by adding DEFINE statements to the code. PROC REPORT uses the variable labels (if not available, the variable names) as column headers for the output. The column headers can be adapted by using DEFINE statements as the following code shows:

```
title 'Defining variables';
proc report data = adsl;
   columns studyid usubjid agegrp sex race ethnic age;
   define studyid / 'Study ID';
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group';
   define sex     / 'Gender';
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age (Years)';
run;
```

The output now displays the labels defined above:

Defining variables 17:09 Wednesday, September 9, 2015 1

Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	11
Test Study	1010003	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	15
Test Study	1010004	Group 2 (6 - < 12 years)	Female	White	Not Hispanic or Latino	9
Test Study	1010005	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	15
Test Study	1010006	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17
Test Study	1010007	Group 1 (12 - < 18 years)	Female	Asian	Not Hispanic or Latino	16
Test Study	1010008	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	16
Test Study	1010009	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	14
Test Study	1010010	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	17

The DEFINE statement offers more options that can be used to change the appearance of the output. The following ones are the most commonly used:

DEFINE OPTIONS - WIDTH

As default PROC REPORT uses the variable length as column width for character variables and 8 for numeric variables, if the variable in the source dataset has no format assigned. In the previous output the column Study ID was wider than needed, whereas the numeric Age column was smaller than needed (causing multiple line breaks in the column header). With the WIDTH option individual column widths can be specified for each variable.

```
title 'Defining variable width';
proc report data = adsl;
   columns studyid usubjid agegrp sex race ethnic age;
   define studyid / 'Study ID'     width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group';
   define sex     / 'Gender';
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age (Years)'  width=11;
run;
```

Submitting this statement, the output looks like this:

Defining variable width 17:09 Wednesday, September 9, 2015 1

Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	11
Test Study	1010003	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	15
Test Study	1010004	Group 2 (6 - < 12 years)	Female	White	Not Hispanic or Latino	9
Test Study	1010005	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	15
Test Study	1010006	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17
Test Study	1010007	Group 1 (12 - < 18 years)	Female	Asian	Not Hispanic or Latino	16
Test Study	1010008	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	16
Test Study	1010009	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	14
Test Study	1010010	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	17

DEFINE OPTIONS - FORMAT

The FORMAT option can be used to assign formats to the variables used in the PROC REPORT. The formats used can either be SAS® formats or a user-defined format.

```
title 'Defining variable format';
proc report data = adsl;
   columns studyid usubjid agegrp sex race ethnic age;
   define studyid / 'Study ID'     width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group';
   define sex     / 'Gender';
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age (Years)'  width=11 format=5.2;
run;
```

The example below shows, how Age is displayed using the 5.2 SAS® format:

Defining variable format 17:09 Wednesday, September 9, 2015 1

Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010003	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	15.00
Test Study	1010004	Group 2 (6 - < 12 years)	Female	White	Not Hispanic or Latino	8.50
Test Study	1010005	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	15.00
Test Study	1010006	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010007	Group 1 (12 - < 18 years)	Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	16.00
Test Study	1010009	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	14.00
Test Study	1010010	Group 1 (12 - < 18 years)	Female	White	Not Hispanic or Latino	17.00

DEFINE OPTIONS - SPACING

The SPACING option in the DEFINE statement specifies how many blank characters should be inserted between columns. The default is 2, but this can be increased or decreased to allow for better readability.

```
title 'Defining spacing';
proc report data = adsl;
   columns studyid usubjid agegrp sex race ethnic age;
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4;
   define sex     / 'Gender'        spacing=4;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age (Years)'   width=11 format=5.2;
run;
```

Now there are 4 instead of previously 2 spaces between the columns:

```
                                    Defining spacing            17:09 Wednesday, September 9, 2015   1

Study ID    Subject ID    Age group                    Gender     Race       Ethnicity                Age (Years)
Test Study  1010001       Group 1 (12 - < 18 years)    Male       White      Not Hispanic or Latino         17.00
Test Study  1010002       Group 2 (6 - < 12 years)     Male       Asian      Not Hispanic or Latino         10.50
Test Study  1010003       Group 1 (12 - < 18 years)    Male       White      Not Hispanic or Latino         15.00
Test Study  1010004       Group 2 (6 - < 12 years)     Female     White      Not Hispanic or Latino          8.50
Test Study  1010005       Group 1 (12 - < 18 years)    Female     White      Not Hispanic or Latino         15.00
Test Study  1010006       Group 1 (12 - < 18 years)    Male       White      Not Hispanic or Latino         17.00
Test Study  1010007       Group 1 (12 - < 18 years)    Female     Asian      Not Hispanic or Latino         16.00
Test Study  1010008       Group 1 (12 - < 18 years)    Male       White      Not Hispanic or Latino         16.00
Test Study  1010009       Group 1 (12 - < 18 years)    Male       White      Not Hispanic or Latino         14.00
Test Study  1010010       Group 1 (12 - < 18 years)    Female     White      Not Hispanic or Latino         17.00
```

DEFINE OPTIONS - ORDER

The ORDER option can be used to order the rows in the report according to their formatted values. The output will display only the first occurrence of the variable with the order option in a set of rows that have the same value for that order variable. The default order is ascending, but this can be changed by adding the DESCENDING option to the DEFINE statement.

```
title 'Defining data order';
proc report data = adsl;
   columns studyid usubjid agegrp sex race ethnic age;
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age (Years)'   width=11 format=5.2;
run;
```

The output below displays the age group in a descending order and with grouped entries:

```
                                    Defining data order         17:09 Wednesday, September 9, 2015   1

Study ID    Subject ID    Age group                    Gender     Race       Ethnicity                Age (Years)
Test Study  1010002       Group 2 (6 - < 12 years)     Male       Asian      Not Hispanic or Latino         10.50
Test Study  1010004                                    Female     White      Not Hispanic or Latino          8.50
Test Study  1010001       Group 1 (12 - < 18 years)    Male       White      Not Hispanic or Latino         17.00
Test Study  1010003                                    Male       White      Not Hispanic or Latino         15.00
Test Study  1010005                                    Female     White      Not Hispanic or Latino         15.00
Test Study  1010006                                    Male       White      Not Hispanic or Latino         17.00
Test Study  1010007                                    Female     Asian      Not Hispanic or Latino         16.00
Test Study  1010008                                    Male       White      Not Hispanic or Latino         16.00
Test Study  1010009                                    Male       White      Not Hispanic or Latino         14.00
Test Study  1010010                                    Female     White      Not Hispanic or Latino         17.00
```

The way the data is ordered can be influenced by specifying further details for the ORDER option. By using "`order=data descending`" values are displayed according to their order in the input data set, whereas "`order=formatted descending`" sorts the data by the formatted values. "`order=internal descending`"

would sort the data in the same way that PROC SORT would sort the data (be aware that this sorting may be platform dependent).

For some statistics it is sometimes helpful to have a sorting variable that should actually not occur in the output. To achieve this the NOPRINT option can additionally be specified in the DEFINE statement of the sorting variable (e.g. "`define sortvar / order=internal noprint`").

DEFINE OPTIONS - JUSTIFICATION

To change the justification of the column entries LEFT, RIGHT or CENTER can be used in the DEFINE statement.

```
title 'Defining data justification';
proc report data = adsl;
   columns studyid usubjid agegrp sex race ethnic age;
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age (Years)'   width=11 format=5.2;
run;
```

Submitting this statement, the entries of gender will be displayed centered:

Defining data justification 17:23 Wednesday, September 9, 2015 1

Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

PROC REPORT LIST OPTION

Specifying all DEFINE statements for a larger output can be annoying and time consuming. Thanks to the LIST option of PROC REPORT the effort needed for this can be reduced. The LIST option allows PROC REPORT to print all statements needed to produce the output into the Log window. Submitting the following code will generate a PROC REPORT with DEFINE statements for all variables in the dataset in the Log window:

```
proc report data = adsl LIST;
run;
```

In the Log window you will find the following code being outputted:

```
NOTE: Multiple concurrent threads will be used to summarize data.
PROC REPORT DATA=WORK.ADSL LS=142 PS=44  SPLIT="/" CENTER ;
COLUMN  DMINFDT SEXD RACED ETHNICD AGE STUDYID AGEGRP SEX RACE ETHNIC VISIT USUBJID;

DEFINE  DMINFDT / SUM FORMAT= DATE9. WIDTH=9     SPACING=2   RIGHT "Date of Sign. on Informed Consent (SAS date)" ;
DEFINE  SEXD / SUM FORMAT= 3. WIDTH=3     SPACING=2   RIGHT "Sex (DVG Number)" ;
DEFINE  RACED / SUM FORMAT= 3. WIDTH=3     SPACING=2   RIGHT "Race (DVG Number)" ;
DEFINE  ETHNICD / SUM FORMAT= 3. WIDTH=3     SPACING=2   RIGHT "Ethnicity (DVG Number)" ;
DEFINE  AGE / SUM FORMAT= 3. WIDTH=3     SPACING=2   RIGHT "Age" ;
DEFINE  STUDYID / DISPLAY FORMAT= $25. WIDTH=25    SPACING=2   LEFT "STUDYID" ;
DEFINE  AGEGRP / DISPLAY FORMAT= $25. WIDTH=25    SPACING=2   LEFT "AGEGRP" ;
DEFINE  SEX / DISPLAY FORMAT= $10. WIDTH=10    SPACING=2   LEFT "SEX" ;
DEFINE  RACE / DISPLAY FORMAT= $10. WIDTH=10    SPACING=2   LEFT "RACE" ;
DEFINE  ETHNIC / DISPLAY FORMAT= $25. WIDTH=25    SPACING=2   LEFT "ETHNIC" ;
DEFINE  VISIT / DISPLAY FORMAT= $15. WIDTH=15    SPACING=2   LEFT "VISIT" ;
DEFINE  USUBJID / DISPLAY FORMAT= $10. WIDTH=10    SPACING=2   LEFT "USUBJID" ;
RUN;

NOTE: There were 10 observations read from the data set WORK.ADSL.
NOTE: The PROCEDURE REPORT printed pages 1-2.
NOTE: PROCEDURE REPORT used (Total process time):
      real time           0.24 seconds
      cpu time            0.02 seconds
```

This code can then serve as a starting point for your adaption by copying it from the Log window into the program window.

SPANNING HEADERS

Sometimes it is necessary to define spanning headers over multiple columns. This is achieved by defining the spanning header in the COLUMNS statement. The code below uses a spanning header across the variables AGEGRP, SEX, RACE, ETHNIC and AGE:

```
title 'Defining spanning headers';
proc report data = adsl;
   columns studyid usubjid ('Demographics' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age (Years)'   width=11 format=5.2;
run;
```

The output produced by this code looks as follows:

Defining spanning headers 17:28 Wednesday, September 9, 2015 1

			Demographics			
Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

Currently it is not clearly visible, that the spanning header flows over the above specified variables. By using a split character a separation line can be introduced to show this.

SPLIT CHARACTER

A split character forces PROC REPORT to create a line break whenever it appears (independent whether they occur in column headers or in the data itself). The character that should cause a line break can be defined via the SPLIT option of the PROC REPORT statement. The code below uses # as split character. A split is done after the spanning header “Demographics” and the following three underscores force PROC REPORT to draw a line in the row that is created by the split character.

In addition the code below uses a split character in the label of the variable AGE (which is defined by the DEFINE statement). The part after the split character “(Years)” is now displayed in a separate line, so that the column header of the variable AGE uses two rows. This allows us to shorten the width that is needed for that column.

```
title 'Using split characters';
proc report data = adsl split='#';
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'   width=7 format=5.2;
run;
```

The output produced by this code looks as follows:

Using split characters 17:28 Wednesday, September 9, 2015 1

		Demographics				
Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

PROC REPORT OPTIONS

The split character shown above is one of many PROC REPORT options, which control how the output that is being produced looks like.
The following examples will explain the most commonly used PROC REPORT options.

HEADLINE AND HEADSKIP

The option HEADLINE draws a line under all column headings and the spaces between them, whereas the HEADSKIP option creates an empty row beneath all column headings (in conjunction with HEADLINE it creates the empty row beneath the separator line of the column headings). The example below shows the usage of these two options:

```
title 'Using the headline and headskip option';
proc report data = adsl split='#' headline headskip;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'   width=7 format=5.2;
run;
```

The output produced is shown below:

Using the headline and headskip option 11:41 Friday, September 11, 2015 1

		Demographics				
Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

SPACING

The SPACING option is available as a DEFINE statement option (as shown before) and also as a general PROC REPORT option. The PROC REPORT SPACING option specifies the number of blank characters between all columns, whereas the DEFINE statement SPACING option specifies it for a single column (overwriting the PROC REPORT SPACING option for single columns).

The PROC REPORT option can be decreased, if more space is needed to show content or it can be increased to improve the readability of the output, as shown below:

```
proc report data = adsl split='#' headline headskip spacing=3;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'   width=7 format=5.2;
run;
```

The output produced by this statement looks as shown below:

Using the spacing option 12:35 Friday, September 11, 2015 1

		Demographics				
Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

NOWINDOW AND MISSING

The PROC REPORT NOWINDOW (or also often written as NOWD) option suppresses the interactive window in SAS® for Windows that allows additional manual edits of the output. This allows the procedure to directly write the output to the specified destination (output window or physical file).

The MISSING option allows PROC REPORT to consider missing values as valid values for group, order, or across variables, which is often very useful in sorted or grouped listings.

SAS® OPTIONS

Despite of the PROC REPORT and DEFINE statements there are also a lot of general SAS® options that control how the produced outputs look like. The most frequently used ones are explained in the subsections below.

PAGESIZE AND LINESIZE

The PAGESIZE and LINESIZE SAS® options control the general width and height of the output being produced. Often fixed values are used for the outputs of specific studies or sponsors. The code below shows how these options can be changed:

```
option pagesize=44 linesize=125;

title 'Using the pagesize and linesize option';
proc report data = adsl split='#' headline headskip spacing=3 nowd missing;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'   width=7 format=5.2;
run;
```

In the output we can see that the line size was increased:

Using the pagesize and linesize option 13:11 Friday, September 11, 2015 1

Demographics

Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

Please note that the LINESIZE and PAGESIZE options (LS and PS are synonyms and can also be used in the code) are also available as PROC REPORT options, which would overwrite the general SAS® options if used. Generally it is recommended to use the general SAS® options as in a lot of companies these are defined in a general standard to keep the output sizes consistent within and across studies.

NOCENTER

The option NOCENTER tells PROC REPORT to left-justify the report and summary text, as shown in the example below:

```
option linesize=125 pagesize=44 nocenter;

title 'Using the nocenter option';
proc report data = adsl split='#' headline headskip spacing=3 nowd missing;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'   width=7 format=5.2;
run;
```

The output being produced by this statement, looks as follows:

Using the nocenter option 13:22 Friday, September 11, 2015 1

Demographics

Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

Please note that the NOCENTER option is also available as PROC REPORT option, which would overwrite the general SAS® options if used.

PAGENO

The SAS® option PAGENO controls the number that is used as starting number for the page numbering. This option is of importance in programs that produce more than one output, as the page number needs to be reset for each output. Otherwise each output would use the last page number of the previous output plus one as starting page number. The example below uses 7 as starting page number:

```
option linesize=125 pagesize=44 nocenter pageno=7;

title 'Using the pageno option';
proc report data = adsl split='#' headline headskip spacing=3 nowd missing;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'     width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'    spacing=4 order descending;
   define sex     / 'Gender'       spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'  width=7 format=5.2;
run;
```

The output being produced by this statement, looks as follows:

Using the pageno option 13:39 Friday, September 11, 2015 7

		Demographics				
Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

NODATE AND NONUMBER

To suppress the output of the date and the page number in the top right corner the SAS® options NODATE and NONUMBER are used. This is often done as the layout needed for the outputs often require the date and page number to be on the last row of each page. The code below shows an example usage of these options:

```
option linesize=125 pagesize=44 nocenter pageno=7 nodate nonumber;

title 'Using the nodate and nonumber option';
proc report data = adsl split='#' headline headskip spacing=3 nowd missing;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'     width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'    spacing=4 order descending;
   define sex     / 'Gender'       spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'  width=7 format=5.2;
run;
```

The output without a date and page number on the top right corner looks as follows:

Using the nodate and nonumber option

Demographics

Study ID	Subject ID	Age group	Gender	Race	Ethnicity	Age (Years)
Test Study	1010002	Group 2 (6 - < 12 years)	Male	Asian	Not Hispanic or Latino	10.50
Test Study	1010004		Female	White	Not Hispanic or Latino	8.50
Test Study	1010001	Group 1 (12 - < 18 years)	Male	White	Not Hispanic or Latino	17.00
Test Study	1010003		Male	White	Not Hispanic or Latino	15.00
Test Study	1010005		Female	White	Not Hispanic or Latino	15.00
Test Study	1010006		Male	White	Not Hispanic or Latino	17.00
Test Study	1010007		Female	Asian	Not Hispanic or Latino	16.00
Test Study	1010008		Male	White	Not Hispanic or Latino	16.00
Test Study	1010009		Male	White	Not Hispanic or Latino	14.00
Test Study	1010010		Female	White	Not Hispanic or Latino	17.00

If the date is needed it could then be retrieved over the system date and time and added as a footnote in the last row, as shown in the following example:

```
option linesize=125 pagesize=44 nocenter pageno=7 nodate nonumber;

title 'Adding the data as a footnote';
footnote "%left(%qsysfunc(date(),date9.))";
proc report data = adsl split='#' headline headskip spacing=3 nowd missing;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'   width=7 format=5.2;
run;
```

To add page numbers please refer to the following paper from Chang Y. Chung, Princeton, NJ and Toby Dunn, Manor, TX: http://changchung.com/download/pageXofY_draft.pdf (accessed on September 11, 2015).

PROC PRINTTO

All previously shown code examples would end up in the output window (in interactive session) or in a file called similar to the program (for batch sessions, e.g. program.out for a program called program.sas). To redirect the output to a physical file named as desired (independent in which type of sessions) PROC PRINTTO needs to be used. An example usage of PROC PRINTTO is shown below:

```
option linesize=125 pagesize=44 nocenter pageno=7 nodate nonumber;
FILENAME outfile "Pathtoyourdirectory/step7e.lst";

PROC PRINTTO FILE=outfile NEW;
RUN;

title 'Adding the data as a footnote';
footnote "%left(%qsysfunc(date(),date9.))";
proc report data = adsl split='#' headline headskip spacing=3 nowd missing;
   columns studyid usubjid ('Demographics#___' (agegrp sex race ethnic age));
   define studyid / 'Study ID'      width=10;
   define usubjid / 'Subject ID';
   define agegrp  / 'Age group'     spacing=4 order descending;
   define sex     / 'Gender'        spacing=4 center;
   define race    / 'Race';
   define ethnic  / 'Ethnicity';
   define age     / 'Age#(Years)'   width=7 format=5.2;
run;

PROC PRINTTO;
RUN;
```

Be aware that there also need to be two PROC PRINTTO calls. One opening call before the PROC REPORT including the filename and one closing call after the PROC REPORT. The NEW option that is used in the opening PROC PRINTTO forces the procedure to clear the old output and to produce the new output after that.

CONCLUSION

The SAS® PROC REPORT procedure is a very powerful procedure.
SAS® provides a very detailed online documentation and there are also quite a few papers online showing tips, tricks and hints for this procedure.

REFERENCES

Base SAS(R) 9.2 Procedures Guide
(http://support.sas.com/documentation/cdl/en/proc/61895/HTML/default/viewer.htm#a002473620.htm accessed on September 11, 2015)

Page X of Y with Proc Report
(http://changchung.com/download/pageXofY_draft.pdf accessed on September 11, 2015)

PROC REPORT Basics: Getting Started with the Primary Statements
(http://support.sas.com/resources/papers/proceedings12/242-2012.pdf accessed on September 11, 2015)

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Nicola Tambascia
Accovion GmbH
Helfmann-Park 10
65760 Eschborn
Work Phone: +49 6196 7709 283
Fax: +49 6196 7709 115
Email: nicola.tambascia@accovion.com
Web: www.accovion.com

Konstanze Morgenroth
Accovion GmbH
Helfmann-Park 10
65760 Eschborn
Work Phone: +49 6196 7709 386
Fax: +49 6196 7709 116
Email: konstanze.morgenroth@accovion.com
Web: www.accovion.com