

So You're Still Not Using PROC REPORT. Why Not?

Ray Pass, inVentiv Health Clinical, St Louis, MO
Daphne Ewing, Gilead Sciences, Inc., Foster City, CA

ABSTRACT

Everyone who can spell SAS knows how to use PROC PRINT. Its primary use may be as a development tool to help in debugging a long multi-step program, or as a simple report generator when all that is really needed is a quick look at the data, or even a basic low-level finished report. However, if a report generation/information delivery tool with powerful formatting, summarizing and analysis features is called for, then PROC REPORT is the solution. PROC REPORT can provide the standard PROC PRINT functionality, but in addition, can easily perform many of the tasks that you would otherwise have to use the SORT, MEANS, FREQ and TABULATE procedures to accomplish. PROC REPORT is part of the Base SAS product, can run in both an interactive screen-painting mode or a batch mode, and should be the basic tool of choice when there is a need to produce powerful and productive reports from SAS data sets. This paper will present the basics of PROC REPORT (non-interactive mode) through a series of progressively more sophisticated examples of code and output.

The original main thrust behind PROC REPORT was its interactive WYSIWYG ability, in which reports can be designed online via screen painting manipulation of actual output. When the desired result is achieved, the design and structure of the report can be saved as a catalog entry and called up repeatedly for additional runs with new sets of data. The design can then be modified at any time, and can either be replaced or saved as an additional entry. This is all done in a Display Manager like environment, and will certainly attract users comfortable with working interactively. The batch language of PROC REPORT did not at first receive the initial acclaim or publicity that was afforded the interactive product. In actual fact however, the batch language is every bit as powerful and feature laden as its online cousin, and is almost a mirror image.

The purpose of this paper is to introduce the batch version of PROC REPORT through a progressive series of annotated examples. It is not meant to be a full introductory text. There have been a series of stand-alone manuals published by the SAS Institute for PROC REPORT, and SAS documentation (hardcopy and online versions) from Release 8 onward includes full coverage of the procedure. In addition, PROC REPORT is fully implemented in the SAS Output Delivery System (ODS) including a very full and powerful use of styles. There are many excellent conference papers in existence as well as SAS documentation covering PROC REPORT in ODS. This basic introduction will not touch upon these features at all.

The data set used throughout this paper is a fictitious survey of the spending habits of 100 college students during an academic year. The source code in **Example-1** describes the data contained in an external file (i.e. C:\PROCREP\PROCREP.DAT). Although the examples in this paper are PC-based, the REPORT procedure is identical on all platforms. The example also contains code to create two SAS formats (\$SEXFMT. and \$YEARFMT.) which are used throughout the paper along with a simple PROC REPORT statement. **Output-1** (see appendix) displays a portion of the PROC REPORT output from this simple run (the full population is not included in any of the non-grouped outputs).

```
data survey;
  infile 'c:\procrep\procrep.dat';

  input @01 fname    $char10.
        @12 linit   $char1.
        @15 sex     $char1.
        @18 acadyear $char1.
        @21 gpa     3.1
        @27 age     2.
        @31 books   4.
        @40 food    4.
        @49 ent     4. ;

run;
*-----;
```

```

proc format;
  value $sexfmt  'M' = 'MALE'
                 'F' = 'FEMALE' ;

  value $yearfmt  '1' = 'FRESHMAN'
                  '2' = 'SOPHOMORE'
                  '3' = 'JUNIOR'
                  '4' = 'SENIOR' ;

run;
*-----;
proc report nowd data=survey;
run;

```

Example-1

As **Output-1** shows, PROC REPORT output in its simplest format is similar to PROC PRINT output. The NOWD option on the REPORT statement is necessary to tell SAS that this is being run in the batch mode, and not in the interactive mode (NOWD is an alias for NOWINDOWS.) **Example-2** produces the report found in **Output-2**, and introduces the basic features of the REPORT Procedure syntax. The first items to note are the HEADLINE and HEADSKIP options on the PROC REPORT statement. These are two of the many formatting and structural options available on this statement, and produce the dashed and blank separator lines between the column headers and data. The COLUMN statement declares the variables to be included in the report structure and the order in which they are to be processed from left to right. Note that there are two variables listed in the COLUMN statement, NAME and TOTAL, which are not found in the input data set, but are rather created for use in the report via the COMPUTE statements described later in the paper. These variables are transient and exist only for the duration of the current REPORT run; they are *not* added to the data set named in the REPORT statement. The series of DEFINE statements is used to assign *optional* formatting information to the variables. The first two, for FNAME and LINIT, instruct REPORT not to print these two variables, via a NOPRINT option. They *must* be included in the COLUMN statement because the NAME variable is computed from the values of these two variables. They are therefore part of the data set being REPORTed, but are not included in the output. The order in which these variables appear in the COLUMN statement is important here, as NAME must be preceded by the variables it is computed from.

```

proc report nowd data=survey headline headskip;

  column fname linit name sex acadyear gpa age books food ent total;

  define fname    / display  noprint;
  define linit    / display  noprint;
  define sex      / display          format=$sexfmt.  width=6;
  define acadyear / display 'ACADEMIC/YEAR' format=$yearfmt. width=9;
  define gpa      / analysis          format=3.1      width=3;
  define age      / analysis          format=3.        width=3;
  define books    / analysis '$-BOOKS'  format=dollar6. width=7;
  define food     / analysis '$-FOOD'   format=dollar6. width=6;
  define ent      / analysis '$-ENTERTAIN' format=dollar6. width=11;
  define name     / computed          width=10;
  define total    / computed '$-TOTAL'  format=dollar6. width=7;

  compute name /char length=10;
    name = trim(fname) || ' ' || linit || '.';
  endcomp;

  compute total;
    total = sum(books.sum, food.sum, ent.sum);
  endcomp;

run;

```

Example-2

PROC REPORT supports six different types of output variables with different uses: DISPLAY, ANALYSIS, COMPUTED, ORDER, GROUP, and ACROSS. In this example, the DEFINE statements declare FNAME, LINIT, SEX and ACADYEAR as DISPLAY variables, GPA, AGE, BOOKS, FOOD and ENT as ANALYSIS variables, and NAME and TOTAL as COMPUTED variables. By default, character variables are DISPLAY and numeric variables are ANALYSIS variables. It is good practice, however, to include these designations anyway; it makes for ease of reading and modification of the code. For DISPLAY and ANALYSIS, all values in the input data set are displayed in

the output. The FORMAT and WIDTH options declare output value formats and column widths, and the quoted strings create the column headers. Note that ACADYEAR has a split header (two header lines separated by a split character). The default split character is the forward slash (“/”), but this can be changed with a SPLIT= option on the REPORT statement if a slash is a valid part of a header or data value. Note also that the WIDTHs must be adequate to contain both the data and the headers (e.g. ‘\$-ENTERTAIN’ needs a WIDTH=11).

The variables NAME and TOTAL were included in the COLUMN statement, and declared in the DEFINE statements as COMPUTED variables. These variables are created in 'compute blocks', each block beginning with a COMPUTE statement and ending with an ENDCOMP statement. The compute blocks contain DATA step statements which create the COMPUTED variables. Options on the COMPUTE statement which create NAME declare it to be a CHARACTER variable with a LENGTH of 10. When referring to an ANALYSIS variable in a compute block, you must use “compound name” syntax (e.g. BOOKS.SUM) to name both the variable and the statistic associated with it, delimiting them with a dot (“.”).

Output-2 displays a useful, formatted, one line per observation report as created with the code in **Example-2**. Much can be done however in terms of ordering the data for presentation. **Example-3** contains the code which produces the categorized and sorted data displayed in **Output-3**. Here the data is ordered by ACADYEAR and SEX.

The first change to note from the code in the previous listing is that the order of variables in the COLUMN statement has been changed. Variables used to order the data, referred to as ORDER variables, must be processed before (to the left of) any DISPLAY variables. This makes intuitive sense as you look at the output, and is backed up by rigid syntactical placement rules. The DEFINE statements declare ACADYEAR and SEX to be ORDER variables. This is what actually creates the sorted display. Note that only the first unique value of each of these variables is displayed. Using ORDER variables eliminates the need for a prior sorting of the data set with PROC SORT. Note that although the sequenced order of variables is critical in the COLUMN statement, it is of no concern in the series of DEFINE statements. The DEFINE statement for ACADYEAR also includes the ORDER= option, which in this case is assigned as INTERNAL. This instructs the REPORT procedure to sort (and display) the observations by their INTERNAL values (1, 2, 3, 4) rather than their formatted values (FRESHMAN, JUNIOR, SENIOR, SOPHOMORE). Other possible presentation orders (FORMATTED, DATA, FREQUENCY) are available as they are in the MEANS, FREQ, and TABULATE procedures. NOTE: The default display order for PROC REPORT is FORMATTED, unlike all the other SAS procedures which use INTERNAL as the default.

```
proc report nowd data=survey headline headskip;

    column acadyear sex fname linit name gpa age books food ent total;

    define fname      / display  noprint;
    define linit      / display  noprint;
    define sex        / order          format=$sexfmt.   width=6;
    define acadyear   / order    'ACADEMIC/YEAR' format=$yearfmt. width=9
                        order=internal;

    define gpa        / analysis      format=3.1       width=3;
    define age        / analysis      format=3.         width=3;
    define books      / analysis    '$-BOOKS'    format=dollar6.  width=7;
    define food       / analysis    '$-FOOD'     format=dollar6.  width=6;
    define ent        / analysis    '$-ENTERTAIN' format=dollar6.  width=11;
    define name       / computed                    width=10;
    define total      / computed    '$-TOTAL'    format=dollar6.  width=7;

    compute name / char length=10;
        name = trim(fname) || ' ' || linit || '.';
    endcomp;

    compute total;
        total = sum(books.sum, food.sum, ent.sum);
    endcomp;

run;
```

Example-3

The next feature of PROC REPORT to be discussed is the ability to collapse (group) observations sharing the same value of any variable or set of variables, and to display summary statistics for each subset of observations. This is

done in **Example-4** by changing ACADYEAR and SEX to GROUP variables, with the results displayed in **Output-4**.

```
proc report nowd data=survey headline headskip;

    column acadyear sex gpa age books food ent total;

    define sex      / group          format=$sexfmt.  width=6;
    define acadyear / group      'ACADEMIC/YEAR' format=$yearfmt. width=9
                                order=internal;

    define gpa      / analysis      format=3.1      width=3;
    define age      / analysis      format=3.        width=3;
    define books    / analysis      '$-BOOKS'        format=dollar6. width=7;
    define food     / analysis      '$-FOOD'          format=dollar7. width=7;
    define ent      / analysis      '$-ENTERTAIN'     format=dollar7. width=11;
    define total    / computed      '$-TOTAL'        format=dollar7. width=7;

    compute total;
        total = sum(books.sum, food.sum, ent.sum);
    endcomp;
run;
```

Example-4

FNAME, LINIT and NAME can not be meaningfully summarized within other (GROUP) variables, and are therefore removed from the COLUMN and DEFINE statements, and also from the summarized report output. Although they can be changed to ORDER or GROUP variables, they can not remain in the design as DISPLAY variables with GROUP variables to their left; and changing them to ORDER or GROUP variables in this example would not provide useful information. Note that the NAME compute block must also be removed. The remaining variables, including the COMPUTED variable TOTAL, are summarized over each ACADYEAR-SEX combination. The resulting sum is displayed in each cell of the report.

Although the SUM statistic might carry some meaning with the financial variables (BOOKS, FOOD, ENT, and TOTAL), a MEAN value would probably be more desirable. For GPA and AGE however, the SUM statistic is meaningless and a MEAN value is absolutely needed. **Example-5** accomplishes these goals by using the 'alias' feature of the REPORT procedure as well as demonstrating the use of different statistics for ANALYSIS variables. This example also demonstrates the powerful ability of the procedure to produce summary break lines, as well as compound column-spanning headers. A few more goodies are demonstrated in the example as well. The resulting much enhanced display is found in **Output-5**.

```
title1 'PROC REPORT - SAMPLE SUMMARY REPORT';
title2 '----- ACADEMIC SURVEY -----';
title3;
*-----;
proc report nowd data=survey split='\';

    column acadyear sex gpa age ('-ANNUAL AMOUNT SPENT ON-'
                                ('-BOOKS-'      books books=booksm)
                                ('-FOOD-'       food  food =foodm)
                                ('-ENTERTAIN-'  ent   ent  =entm)
                                ('-TOTAL-'      total totalm));

    define sex      / group      'SEX\--'          format=$sexfmt.  width=6;
    define acadyear / group      'ACADEMIC\YEAR\--'    format=$yearfmt. width=9
                                order=internal;

    define gpa      / analysis    'MEAN\GPA\--'        format=3.1 mean  width=4;
    define age      / analysis    'MEAN\AGE\--'        format=3. mean   width=4;
    define books    / analysis    'TOTAL\--'          format=dollar7. width=7  sum;
    define booksm   / analysis    'MEAN\--'           format=dollar4. width=4  mean;
    define food     / analysis    'TOTAL\--'          format=dollar7. width=7  sum;
    define foodm    / analysis    'MEAN\--'           format=dollar4. width=4  mean;
    define ent      / analysis    'TOTAL\--'          format=dollar7. width=7  sum;
    define entm     / analysis    'MEAN\--'           format=dollar4. width=4  mean;
```

```

define total      / computed 'TOTAL\--'          format=dollar8.  width=8;
define totalm    / computed 'MEAN\--'          format=dollar7.  width=7;

compute total;
  total = sum(books.sum, food.sum, ent.sum);
endcomp;

compute totalm;
  totalm = sum(bookm, foodm, entm);
endcomp;

break after acadyear / ol ul skip summarize suppress;

rbreak after / dol dul skip summarize;

run;

```

Example-5

Much is happening in this example. The HEADLINE and HEADSKIP options are removed from the REPORT statement in lieu of another method of creating header underlines. Each DEFINE statement now contains an additional final segment in its title option, namely '\--'. These two dashes will automatically expand to the width of the column, yielding a broken, but perhaps more pleasing set of column header underlines (the keyword is 'perhaps' as it is somewhat subjective, but at least the alternative is available). Notice that the SPLIT='\ ' option is used (any character can be defined as a line splitting delimiter). When this technique is used, all variables should be given this type of title for uniformity of display.

The next major change is the enhanced COLUMN statement. Aliases are created for the BOOKS, FOOD and ENT variables, namely BOOKM, FOODM and ENTM. This is done so that the information for each of these variables can be presented in two different manners via separate DEFINE statements for each variable and its alias. Each variable is displayed first as a SUM (this is the default presentation for all ANALYSIS variables) and then as a MEAN, by including these statistics in the appropriate DEFINE statements. New variables are *not* being created; each variable with an alias is merely being displayed multiple times.

The COLUMN statement now also contains column-spanning headers. These can be simple (a header which spans multiple columns) or compound as in this example. A column spanning header is created by including a quoted string in parentheses preceding the columns it is to span. In the header, the initial and final characters can be extended to the total width of the set of columns it spans. In our example, the uppermost header, '-ANNUAL AMOUNT SPENT ON-' spans all the displayed variables, and the next level down headers, '-BOOKS-', etc. each span one variable and its alias.

The DEFINE statements for the ANALYSIS variables and their aliases, and for the COMPUTED variables TOTAL and TOTALM are enhanced in various ways. REPORT is now instructed to display each of the GROUP and ANALYSIS variables as both a MEAN and a SUM (if no statistic is presented, the default SUM is used). In addition, WIDTHS and FORMATS are changed judiciously (trial and error) to accommodate the new headers and data displayed. A new COMPUTED variable TOTALM is now calculated from the aliases of the components making up TOTAL. Since they are MEAN values, the new variable TOTALM will reflect MEAN values as well. Note that aliases do *not* use the compound dot notation.

Another enhancement to the new display is the addition of summary break lines to the report. The BREAK command inserts a break line AFTER each new value of ACADYEAR (the keyword BEFORE can also be used to place break lines before grouped variables). Each new line has a summary statistic (SUMMARIZE) with an overline (OL), an underline (UL) and a blank line (SKIP) before the next printed line. The RBREAK command instructs REPORT to print a break line AFTER the entire report (BEFORE is also available), with double over and underlining (DOL, DUL) and overall summary statistics (SUMMARIZE). Each of the BREAK commands also uses the SUPPRESS option to prevent printing BREAK information (values, under- and overlining) under the GROUP variable columns (ACADYEAR, SEX). The last enhancements to the code in **Example-5** is the addition of TITLE statements. These statements are standard SAS syntax, and are not saved with the REPORT definition if it is stored as a catalog entry.

Output-5 presents a useful display of grouped information. PROC REPORT also allows other types of useful formatted output, both grouped and non-grouped. **Example-6** contains code used to display the telephone-book-column style display found in **Output-6**.

In **Example-6**, the REPORT statement contains three new options which instruct the procedure to display the data in three repeating 'snaking' sets of variables (PANELS) with a pagesize (PS) of 50 lines, and 10 spaces between each panel (PSPACE). Only NAME, ACADYEAR and AGE are printed (note that FNAME and LINIT must be included in the COLUMN statement for processing, but are not printed via NOPRINT).

```

title1 'PROC REPORT - SAMPLE COLUMNAR REPORT';
title2 '----- ACADEMIC SURVEY -----';
title3;
*-----;
proc report nowd data=survey headline headskip panels=3 ps=50 pspace=10;

  column fname linit name acadyear gpa;

  define fname    / display noprint;
  define linit    / display noprint;
  define acadyear / display 'ACADEMIC/YEAR' format=$yearfmt. width=9;
  define gpa      / analysis          format=3.1          width=3;
  define name     / computed          width=10;

  compute name / char length=10;
    name = trim(fname)||' '||linit||'.';
  endcomp;
run;

```

Example-6

Example-7 also uses PANELS to display its data, but now uses the FLOW option to achieve a different presentation style. The FLOW option instructs the procedure to wrap the value of a character variable within its column width. In this example, the COMPUTED character variable BLOCK is constructed to include the defined split character '\ between each of its components. As the FLOW option honors split characters, the effect is a wrapping to a new line for each component in BLOCK. The FLOW option can easily be used to generate name and address listings. Note that all the components that make up BLOCK are required in the COLUMN statement, but they are all NOPRINTED. Note also that an artificially large number of PANELS (99) is stated. This technique allows the procedure to determine how many PANELS to produce based on the linesize used.

```

title1 'PROC REPORT - SAMPLE FLOW REPORT';
title2 '----- ACADEMIC SURVEY -----';
title3;
*-----;
proc report nowd data=survey split='\ ' panels=99 ps=50 ;

  column fname linit acadyear age sex gpa block;

  define  fname    / display noprint;
  define  linit    / display noprint;
  define  acadyear / display noprint;
  define  age      / display noprint;
  define  sex      / display noprint;
  define  gpa      / display noprint;
  define  block    / computed          flow width=20 ' ';

  compute block / char length=100;
    block = 'name: '||trim(fname)||' '||linit||'.'
           ||'\YEAR: '||put(acadyear,$yearfmt.)
           ||'\ AGE: '||put(age,2.)
           ||'\ SEX: '||put(sex,$sexfmt.)
           ||'\ GPA: '||put(gpa,3.1)
           ||'\ \';
  endcomp;
run;

```

Example-7

Example-8 also contains code for a non-grouped data display as shown in **Output-8**. In this example, the WRAP option in the REPORT statement causes the values for all variables in an observation to be printed in a line-wrap fashion before printing of the next observation is begun.

```

title1 'PROC REPORT - SAMPLE WRAPPED REPORT';
title2 '----- ACADEMIC SURVEY -----';
title3;
*-----;
data survey2;
  set survey;
  obsno = _n_;
run;
*-----;
proc report nowd data=survey2 headline headskip named wrap ls=70;

  column obsno fname linit name acadyear age sex gpa books food ent total;

  define obsno    / order    noprint;
  define fname    / display  noprint;
  define linit    / display  noprint;
  define sex      / display          format=$sexfmt.   width=6;
  define acadyear / display  'YEAR'  format=$yearfmt. width=9;
  define gpa      / analysis        format=3.1         width=3;
  define age      / analysis        format=3.          width=3;
  define books    / analysis  '$-BOOKS' format=dollar6.  width=6;
  define food     / analysis  '$-FOOD'  format=dollar6.  width=6 spacing=3;
  define ent      / analysis  '$-ENTERTAIN' format=dollar6.  width=6 spacing=3;
  define name     / computed                          width=10;
  define total    / computed  '$-TOTAL'  format=dollar6.  width=7 spacing=3;

  compute name /char length=10;
    name = trim(fname)||' '||linit||'.';
  endcomp;

  compute total;
    total = sum(books.sum, food.sum, ent.sum);
  endcomp;

  break after obsno / skip;
run;

```

Example-8

The NAMED option causes each variable to be preceded by its header (variable names would be used if no headers were presented) followed by “=”. The LS or Linesize option causes REPORT to restrict horizontal printing to 70 columns. Some of the DEFINE statements use a SPACING option to manipulate the number of blank spaces printed before each value. This was done to create the quasi columnar alignment which makes for easier reading. Each observation is also followed by a blank line before the next one is started. The REPORT procedure does not at present provide for double-spacing, so it had to be tricked into it. This was done with the aid of some pre-processing of the data set before the REPORT procedure was called. Data set SURVEY2 was created by SET'ing SURVEY, and then creating a new variable, OBSNO, by assigning it the value of the observation counter (_N_), an automatic system variable provided by the SAS system. The REPORT procedure was then run against data set SURVEY2. In the REPORT code, OBSNO is DEFINE'd as an ORDER variable, and a BREAK statement is used with the SKIP option to add a blank line after each new value of OBSNO. As each value of OBSNO (_N_) is unique, the effect is to skip a line after each new observation printed.

The final example is shown in **Output-9**, and was created with the code found in **Example-9**. This example demonstrates the last variable type allowed in PROC REPORT. The variables included in the example are ACADYEAR, SEX, GPA, where SEX is DEFINE'd as an ACROSS variable. This causes its values to be grouped like a GROUP variable, but displayed horizontally rather than vertically. The next enhancement is in the format of the COLUMN statement. The order of processing the variable is: ACADYEAR SEX GPA. The “,” in the statement after SEX instructs REPORT to print the summary values of the ANALYSIS variable GPA *under* the unique values of the ACROSS variable SEX (if no statistics were included for GPA, the frequency of observations in each cell

would be printed). This syntax allows PROC REPORT to display its data in a simple PROC TABULATE like manner, and demonstrates more of the power of the REPORT procedure.

```

title1 'PROC REPORT - SAMPLE ACROSS REPORT';
title2 '----- ACADEMIC SURVEY -----';
title3;
*-----;
proc report nowd data=survey headline headskip out=out_ex;

    column acadyear sex,gpa;

    define sex      / across          format=$sexfmt.  width=6;
    define acadyear / group      'ACADEMIC/YEAR' format=$yearfmt. width=9;
    define gpa      / analysis      format=3.1        width=6   mean;
run;
*-----;
title1 "CONTENTS OF PROC REPORT OUTPUT DATA SET";
title2 "----- PRODUCED WITH 'OUT=' OPTION -----";
title3;
*-----;
proc print data=out_ex;
run;

```

Example-9

The final addition in **Example-9** is the use of the output data set option, `OUT=output data set`, in the REPORT statement. This instructs the procedure to produce an output SAS data set containing one observation for each output line of data in the REPORT output. If the report contains summary (BREAK) lines, then the name of the BREAK variable is contained in a new variable called “_BREAK_”. If there are no BREAKs, the value of “_BREAK_” is missing. We include a simple PROC PRINT here (we are not REPORT bigots) to display the contents of the output data set OUT_EX. Whenever an ACROSS variable is used, PROC REPORT uses a “_Cn_” notation to name each column. In this example, the two values of the ACROSS variable SEX (“FEMALE” and “MALE”) are represented in OUT_EX as _C2_ and _C3_ respectively. Getting used to the “_Cn_” notation is a chore, but well worth it. Although other procedures can produce output data sets of summary values, using ACROSS variables with output data sets can be used to produce summarized cross-tabular output data.

There are many additional features in the REPORT procedure which have not been included in this paper due to the desire to keep this on a very basic introductory level, but which add significantly to the utility of the procedure. One example is the ability to include strings of literal text in the body of the output, and the ability to conditionally include whole segments of output based on values of data set or COMPUTED variables. Other new features are being developed and will be included in future releases of the SAS System base product, and its accompanying documentation.

CONCLUSION

The REPORT procedure is a powerful and productive tool in the SAS user's report writing arsenal, and should be fully examined and exploited. It provides a much expanded realm of utility over the PRINT procedure alone. PROC REPORT can do much of the work of PROC SORT, PROC MEANS and a fair amount of PROC TABULATE. It is by no means a reporting panacea, and should never be used to the total exclusion of other procedures. It can, and should however be used appropriately and cooperatively with other SAS System tools to provide a powerful synergy.

So if you're still not using PROC REPORT, or have no intention to at least investigate it, once again we can only ask, *Why Not?*

REFERENCES

SAS is a registered trademark of the SAS Institute Inc., Cary, NC, USA.

AUTHORS

The authors of this paper can be contacted as follows:

Ray Pass, inVentiv Health Clinical
Voice: (314) 202-0208
E-mail: ray.pass@inventivhealth.com

Daphne Ewing, Gilead Sciences, Inc.
Voice: (215) 740-6503
E-mail: DEwing@comcast.net

The SAS System

FNAME	L	I	N	S	E	GPA	AGE	BOOKS	FOOD	ENT
DORY	A	F	3			4.0	20	381	945	1144
GERALD	A	M	3			3.5	20	265	655	687
HARRY	A	M	4			3.6	19	544	855	478
REGGIE	A	F	3			3.6	20	100	841	654
VALERIE	A	F	3			3.2	20	532	756	250
ANTHONY	B	M	1			2.0	16	425	855	965
GEORGE	B	M	3			2.5	21	522	752	425
TOM	B	M	2			3.8	19	215	625	622
CAROL	C	F	4			1.6	21	344	1200	602
EDWARD	C	M	1			2.9	18	451	466	2101
KATHY	C	F	4			2.5	21	544	842	874
BILL	D	M	3			3.8	18	244	810	620
CARMEN	D	F	4			3.9	18	522	1022	587
DONNA	D	F	1			3.2	19	652	802	785
FRANK	D	M	4			3.6	21	675	841	254
JOAN	D	F	2			2.0	27	680	789	855
LEON	D	M	1			2.9	18	201	755	400
ROSE	D	F	2			3.6	18	205	866	100
TOM	D	M	2			3.6	18	406	485	544
ALLAN	E	M	3			3.5	21	345	455	852
BETTY	E	F	2			3.6	20	222	755	962
DENNIS	E	M	3			2.8	19	541	480	221
EVELYN	E	F	2			3.2	19	265	855	452

Output-1

The SAS System

NAME	SEX	ACADEMIC YEAR	GPA	AGE	\$-BOOKS	\$-FOOD	\$-ENTERTAIN	\$-TOTAL
DORY A.	FEMALE	JUNIOR	4.0	20	\$381	\$945	\$1,144	\$2,470
GERALD A.	MALE	JUNIOR	3.5	20	\$265	\$655	\$687	\$1,607
HARRY A.	MALE	SENIOR	3.6	19	\$544	\$855	\$478	\$1,877
REGGIE A.	FEMALE	JUNIOR	3.6	20	\$100	\$841	\$654	\$1,595
VALERIE A.	FEMALE	JUNIOR	3.2	20	\$532	\$756	\$250	\$1,538
ANTHONY B.	MALE	FRESHMAN	2.0	16	\$425	\$855	\$965	\$2,245
GEORGE B.	MALE	JUNIOR	2.5	21	\$522	\$752	\$425	\$1,699
TOM B.	MALE	SOPHOMORE	3.8	19	\$215	\$625	\$622	\$1,462
CAROL C.	FEMALE	SENIOR	1.6	21	\$344	\$1,200	\$602	\$2,146
EDWARD C.	MALE	FRESHMAN	2.9	18	\$451	\$466	\$2,101	\$3,018
KATHY C.	FEMALE	SENIOR	2.5	21	\$544	\$842	\$874	\$2,260
BILL D.	MALE	JUNIOR	3.8	18	\$244	\$810	\$620	\$1,674
CARMEN D.	FEMALE	SENIOR	3.9	18	\$522	\$1,022	\$587	\$2,131
DONNA D.	FEMALE	FRESHMAN	3.2	19	\$652	\$802	\$785	\$2,239
FRANK D.	MALE	SENIOR	3.6	21	\$675	\$841	\$254	\$1,770
JOAN D.	FEMALE	SOPHOMORE	2.0	27	\$680	\$789	\$855	\$2,324
LEON D.	MALE	FRESHMAN	2.9	18	\$201	\$755	\$400	\$1,356
ROSE D.	FEMALE	SOPHOMORE	3.6	18	\$205	\$866	\$100	\$1,171
TOM D.	MALE	SOPHOMORE	3.6	18	\$406	\$485	\$544	\$1,435
ALLAN E.	MALE	JUNIOR	3.5	21	\$345	\$455	\$852	\$1,652
BETTY E.	FEMALE	SOPHOMORE	3.6	20	\$222	\$755	\$962	\$1,939
DENNIS E.	MALE	JUNIOR	2.8	19	\$541	\$480	\$221	\$1,242
EVELYN E.	FEMALE	SOPHOMORE	3.2	19	\$265	\$855	\$452	\$1,572
JONAH E.	MALE	SENIOR	3.2	17	\$569	\$785	\$457	\$1,811

Output-2

The SAS System								
ACADEMIC YEAR	SEX	NAME	GPA	AGE	\$-BOOKS	\$-FOOD	\$-ENTERTAIN	\$-TOTAL

FRESHMAN	FEMALE	DONNA D.	3.2	19	\$652	\$802	\$785	\$2,239
		LILLY E.	3.6	18	\$489	\$966	\$954	\$2,409
		JANE K.	2.9	18	\$402	\$1,300	\$541	\$2,243
		JACKIE M.	2.1	16	\$402	\$1,200	\$454	\$2,056
		JOAN O.	2.8	18	\$306	\$522	\$777	\$1,605
		DEBBIE R.	3.8	18	\$432	\$230	\$556	\$1,218
		SALLY R.	3.8	18	\$524	\$600	\$400	\$1,524
		ERICA S.	4.0	18	\$560	\$455	\$201	\$1,216
		ARLENE T.	3.5	18	\$512	\$466	\$412	\$1,390
		ANTHONY B.	2.0	16	\$425	\$855	\$965	\$2,245
	MALE	EDWARD C.	2.9	18	\$451	\$466	\$2,101	\$3,018
		LEON D.	2.9	18	\$201	\$755	\$400	\$1,356
		SAM E.	3.4	18	\$894	\$688	\$865	\$2,447
		JOHN F.	3.8	20	\$512	\$741	\$905	\$2,158
		JOEL H.	2.5	24	\$524	\$987	\$950	\$2,461
		MIKE J.	3.2	19	\$254	\$522	\$656	\$1,432
		FRANK L.	3.5	19	\$478	\$946	\$125	\$1,549
		GEORGE L.	2.8	18	\$490	\$954	\$795	\$2,239
		JIM R.	3.7	38	\$452	\$866	\$555	\$1,873
		ROBERT S.	3.8	17	\$500	\$547	\$220	\$1,267
SOPHOMORE	FEMALE	MANNY V.	3.4	34	\$325	\$862	\$544	\$1,731
		JOAN D.	2.0	27	\$680	\$789	\$855	\$2,324
		ROSE D.	3.6	18	\$205	\$866	\$100	\$1,171
		BETTY E.	3.6	20	\$222	\$755	\$962	\$1,939
		EVELYN E.	3.2	19	\$265	\$855	\$452	\$1,572
		CHERYL F.	3.8	19	\$245	\$577	\$622	\$1,444
		CINDY F.	3.7	19	\$318	\$622	\$455	\$1,395
		LAURIE F.	3.2	19	\$345	\$844	\$650	\$1,839
		CONNIE K.	2.9	19	\$156	\$544	\$999	\$1,699
		MARIE K.	3.2	19	\$523	\$684	\$787	\$1,994
		ELLEN T.	3.0	19	\$241	\$622	\$956	\$1,819
		MILLY T.	4.0	28	\$325	\$876	\$557	\$1,758
		ROBIN Y.	3.5	20	\$345	\$788	\$454	\$1,587
		DINA Z.	3.5	18	\$255	\$655	\$445	\$1,355
		MALE	TOM B.	3.8	19	\$215	\$625	\$622
TOM D.	3.6		18	\$406	\$485	\$544	\$1,435	

Output-3

The SAS System							
ACADEMIC YEAR	SEX	GPA	AGE	\$-BOOKS	\$-FOOD	\$-ENTERTAIN	\$-TOTAL

FRESHMAN	FEMALE	30	161	\$4,279	\$6,541	\$5,080	\$15,900
	MALE	38	259	\$5,506	\$9,189	\$9,081	\$23,776
SOPHOMORE	FEMALE	43	264	\$4,125	\$9,477	\$8,294	\$21,896
	MALE	57	335	\$5,115	\$12,034	\$10,979	\$28,128
JUNIOR	FEMALE	45	285	\$5,985	\$10,820	\$8,122	\$24,927
	MALE	61	393	\$7,185	\$15,963	\$9,860	\$33,008
SENIOR	FEMALE	21	143	\$3,031	\$6,281	\$4,696	\$14,008
	MALE	26	207	\$5,164	\$6,801	\$5,007	\$16,972

Output-4

PROC REPORT - SAMPLE SUMMARY REPORT
 ----- ACADEMIC SURVEY -----

ACADEMIC YEAR	SEX	MEAN GPA	MEAN AGE	-----ANNUAL AMOUNT SPENT ON-----							
				----BOOKS----		----FOOD----		--ENTERTAIN--		-----TOTAL-----	
				TOTAL	MEAN	TOTAL	MEAN	TOTAL	MEAN	TOTAL	MEAN
FRESHMAN	FEMALE	3.3	18	\$4,279	\$475	\$6,541	\$727	\$5,080	\$564	\$15,900	\$1,767
	MALE	3.2	22	\$5,506	\$459	\$9,189	\$766	\$9,081	\$757	\$23,776	\$1,981
		3.2	20	\$9,785	\$466	\$15,730	\$749	\$14,161	\$674	\$39,676	\$1,889
SOPHOMORE	FEMALE	3.3	20	\$4,125	\$317	\$9,477	\$729	\$8,294	\$638	\$21,896	\$1,684
	MALE	3.3	20	\$5,115	\$301	\$12,034	\$708	\$10,979	\$646	\$28,128	\$1,655
		3.3	20	\$9,240	\$308	\$21,511	\$717	\$19,273	\$642	\$50,024	\$1,667
JUNIOR	FEMALE	3.2	20	\$5,985	\$428	\$10,820	\$773	\$8,122	\$580	\$24,927	\$1,781
	MALE	3.2	21	\$7,185	\$378	\$15,963	\$840	\$9,860	\$519	\$33,008	\$1,737
		3.2	21	\$13,170	\$399	\$26,783	\$812	\$17,982	\$545	\$57,935	\$1,756
SENIOR	FEMALE	3.0	20	\$3,031	\$433	\$6,281	\$897	\$4,696	\$671	\$14,008	\$2,001
	MALE	2.9	23	\$5,164	\$574	\$6,801	\$756	\$5,007	\$556	\$16,972	\$1,886
		2.0	22	\$8,195	\$512	\$13,082	\$818	\$9,703	\$606	\$30,980	\$1,936
		3.2	20	\$40,390	\$404	\$77,106	\$771	\$61,119	\$611	\$178,615	\$1,786

Output-5

PROC REPORT - SAMPLE COLUMNAR REPORT
 ----- ACADEMIC SURVEY -----

NAME	ACADEMIC YEAR	GPA	NAME	ACADEMIC YEAR	GPA	NAME	ACADEMIC YEAR	GPA
DORY A.	JUNIOR	4.0	SALLY I.	JUNIOR	3.5	ARLENE T.	FRESHMAN	3.5
GERALD A.	JUNIOR	3.5	WILLIAM I.	SOPHOMORE	3.6	BEN T.	JUNIOR	3.2
HARRY A.	SENIOR	3.6	ALAN J.	JUNIOR	3.4	ELLEN T.	SOPHOMORE	3.0
REGGIE A.	JUNIOR	3.6	MIKE J.	FRESHMAN	3.2	MILLY T.	SOPHOMORE	4.0
VALERIE A.	JUNIOR	3.2	CONNIE K.	SOPHOMORE	2.9	WILL T.	SENIOR	2.9
ANTHONY B.	FRESHMAN	2.0	DAVID K.	SOPHOMORE	3.8	ANNA V.	JUNIOR	2.3
GEORGE B.	JUNIOR	2.5	EILEEN K.	SENIOR	3.0	MANNY V.	FRESHMAN	3.4
TOM B.	SOPHOMORE	3.8	JANE K.	FRESHMAN	2.9	DAVID W.	JUNIOR	2.5
CAROL C.	SENIOR	1.6	JOSE K.	SOPHOMORE	3.5	PAUL W.	SENIOR	3.0
EDWARD C.	FRESHMAN	2.9	MARIE K.	SOPHOMORE	3.2	AL X.	SOPHOMORE	2.5
KATHY C.	SENIOR	2.5	PETER K.	SOPHOMORE	3.6	ROBIN Y.	SOPHOMORE	3.5
BILL D.	JUNIOR	3.8	TERRY K.	JUNIOR	4.0	SAM Y.	SOPHOMORE	4.0
CARMEN D.	SENIOR	3.9	FRANK L.	FRESHMAN	3.5	TONY Y.	SENIOR	2.0
DONNA D.	FRESHMAN	3.2	GEORGE L.	FRESHMAN	2.8	DINA Z.	SOPHOMORE	3.5
FRANK D.	SENIOR	3.6	SHEILA L.	SENIOR	4.0			
JOAN D.	SOPHOMORE	2.0	STEVE L.	SOPHOMORE	3.5			
LEON D.	FRESHMAN	2.9	DAN M.	JUNIOR	1.6			
ROSE D.	SOPHOMORE	3.6	DANIEL M.	SOPHOMORE	2.3			
TOM D.	SOPHOMORE	3.6	JACKIE M.	FRESHMAN	2.1			
ALLAN E.	JUNIOR	3.5	LOU M.	JUNIOR	2.0			
BETTY E.	SOPHOMORE	3.6	FRED O.	SENIOR	2.5			
DENNIS E.	JUNIOR	2.8	JOAN O.	FRESHMAN	2.8			
EVELYN E.	SOPHOMORE	3.2	HARRIET P.	JUNIOR	3.0			
JONAH E.	SENIOR	3.2	PAUL P.	SOPHOMORE	2.6			
LILLY E.	FRESHMAN	3.6	LARRY Q.	JUNIOR	3.7			
MARIA E.	JUNIOR	3.4	CARL R.	JUNIOR	3.6			
SAM E.	FRESHMAN	3.4	CYNTHIA R.	JUNIOR	3.0			
WILLIAM E.	SOPHOMORE	3.7	DEBBIE R.	FRESHMAN	3.8			
CHERYL F.	SOPHOMORE	3.8	HENRY R.	JUNIOR	3.8			

Output-6

PROC REPORT - SAMPLE FLOW REPORT
 ----- ACADEMIC SURVEY -----

NAME: DORY A. YEAR: JUNIOR AGE: 20 SEX: FEMALE GPA: 4.0	NAME: TOM B. YEAR: SOPHOMORE AGE: 19 SEX: MALE GPA: 3.8	NAME: FRANK D. YEAR: SENIOR AGE: 21 SEX: MALE GPA: 3.6	NAME: DENNIS E. YEAR: JUNIOR AGE: 19 SEX: MALE GPA: 2.8
NAME: GERALD A. YEAR: JUNIOR AGE: 20 SEX: MALE GPA: 3.5	NAME: CAROL C. YEAR: SENIOR AGE: 21 SEX: FEMALE GPA: 1.6	NAME: JOAN D. YEAR: SOPHOMORE AGE: 27 SEX: FEMALE GPA: 2.0	NAME: EVELYN E. YEAR: SOPHOMORE AGE: 19 SEX: FEMALE GPA: 3.2
NAME: HARRY A. YEAR: SENIOR AGE: 19 SEX: MALE GPA: 3.6	NAME: EDWARD C. YEAR: FRESHMAN AGE: 18 SEX: MALE GPA: 2.9	NAME: LEON D. YEAR: FRESHMAN AGE: 18 SEX: MALE GPA: 2.9	NAME: JONAH E. YEAR: SENIOR AGE: 17 SEX: MALE GPA: 3.2
NAME: REGGIE A. YEAR: JUNIOR AGE: 20 SEX: FEMALE GPA: 3.6	NAME: KATHY C. YEAR: SENIOR AGE: 21 SEX: FEMALE GPA: 2.5	NAME: ROSE D. YEAR: SOPHOMORE AGE: 18 SEX: FEMALE GPA: 3.6	NAME: LILLY E. YEAR: FRESHMAN AGE: 18 SEX: FEMALE GPA: 3.6
NAME: VALERIE A. YEAR: JUNIOR AGE: 20 SEX: FEMALE GPA: 3.2	NAME: BILL D. YEAR: JUNIOR AGE: 18 SEX: MALE GPA: 3.8	NAME: TOM D. YEAR: SOPHOMORE AGE: 18 SEX: MALE GPA: 3.6	NAME: MARIA E. YEAR: JUNIOR AGE: 20 SEX: FEMALE GPA: 3.4

Output-7

PROC REPORT - SAMPLE WRAPPED REPORT
 ----- ACADEMIC SURVEY -----

```

-----
NAME=DORY A.      YEAR=JUNIOR      AGE= 20  SEX=FEMALE  GPA=4.0
$-BOOKS= $381   $-FOOD= $945   $-ENTERTAIN=$1,144  $-TOTAL= $2,470

NAME=GERALD A.   YEAR=JUNIOR      AGE= 20  SEX=MALE   GPA=3.5
$-BOOKS= $265   $-FOOD= $655   $-ENTERTAIN= $687  $-TOTAL= $1,607

NAME=HARRY A.    YEAR=SENIOR      AGE= 19  SEX=MALE   GPA=3.6
$-BOOKS= $544   $-FOOD= $855   $-ENTERTAIN= $478  $-TOTAL= $1,877

NAME=REGGIE A.   YEAR=JUNIOR      AGE= 20  SEX=FEMALE  GPA=3.6
$-BOOKS= $100   $-FOOD= $841   $-ENTERTAIN= $654  $-TOTAL= $1,595

NAME=VALERIE A.  YEAR=JUNIOR      AGE= 20  SEX=FEMALE  GPA=3.2
$-BOOKS= $532   $-FOOD= $756   $-ENTERTAIN= $250  $-TOTAL= $1,538

NAME=ANTHONY B.  YEAR=FRESHMAN    AGE= 16  SEX=MALE   GPA=2.0
$-BOOKS= $425   $-FOOD= $855   $-ENTERTAIN= $965  $-TOTAL= $2,245

NAME=GEORGE B.   YEAR=JUNIOR      AGE= 21  SEX=MALE   GPA=2.5
$-BOOKS= $522   $-FOOD= $752   $-ENTERTAIN= $425  $-TOTAL= $1,699

NAME=TOM B.      YEAR=SOPHOMORE   AGE= 19  SEX=MALE   GPA=3.8
$-BOOKS= $215   $-FOOD= $625   $-ENTERTAIN= $622  $-TOTAL= $1,462

NAME=CAROL C.    YEAR=SENIOR      AGE= 21  SEX=FEMALE  GPA=1.6
$-BOOKS= $344   $-FOOD=$1,200  $-ENTERTAIN= $602  $-TOTAL= $2,146

NAME=EDWARD C.   YEAR=FRESHMAN    AGE= 18  SEX=MALE   GPA=2.9
$-BOOKS= $451   $-FOOD= $466   $-ENTERTAIN=$2,101  $-TOTAL= $3,018

NAME=KATHY C.    YEAR=SENIOR      AGE= 21  SEX=FEMALE  GPA=2.5
$-BOOKS= $544   $-FOOD= $842   $-ENTERTAIN= $874  $-TOTAL= $2,260

NAME=BILL D.     YEAR=JUNIOR      AGE= 18  SEX=MALE   GPA=3.8
$-BOOKS= $244   $-FOOD= $810   $-ENTERTAIN= $620  $-TOTAL= $1,674

NAME=CARMEN D.   YEAR=SENIOR      AGE= 18  SEX=FEMALE  GPA=3.9
$-BOOKS= $522   $-FOOD=$1,022  $-ENTERTAIN= $587  $-TOTAL= $2,131

NAME=DONNA D.    YEAR=FRESHMAN    AGE= 19  SEX=FEMALE  GPA=3.2
$-BOOKS= $652   $-FOOD= $802   $-ENTERTAIN= $785  $-TOTAL= $2,239

NAME=FRANK D.    YEAR=SENIOR      AGE= 21  SEX=MALE   GPA=3.6
$-BOOKS= $675   $-FOOD= $841   $-ENTERTAIN= $254  $-TOTAL= $1,770
  
```

Output-8

PROC REPORT-SAMPLE ACROSS REPORT
 ----- ACADEMIC SURVEY -----

ACADEMIC YEAR	SEX	
	FEMALE GPA	MALE GPA
FRESHMAN	3.3	3.2
SOPHOMORE	3.3	3.3
JUNIOR	3.2	3.2
SENIOR	3.0	2.9

Output-9a

CONTENTS OF PROC REPORT OUTPUT DATA SET
 ----- PRODUCED WITH 'OUT=' OPTION -----

OBS	ACADYEAR	_C2_	_C3_	_BREAK_
1	1	3.30000	3.15833	
2	2	3.32308	3.34706	
3	3	3.24286	3.22632	
4	4	3.00000	2.90000	

Output-9b