

Procedures

Why do I get missing values for all rows in a COMPUTED column in PROC REPORT?

The REPORT Procedure is a veritable Swiss Army knife of a procedure and can turn its hand to many different forms of reporting, but with this flexibility comes the occasional oddity which may stump you.

The COLUMN statement in PROC REPORT serves two purposes: it determines which are the columns of interest to the procedure, and the order in which they will appear in the finished report. Consider the code:

```
proc report data = sashelp.shoes nowd ;
  where region = 'Western Europe' ;
  column subsidiary product sales returns inventory ;
  define subsidiary / order
;
  define product      / display
;
  define pctsold      / computed format = percent8.2  'Percentage/Sold'
;
  define sales        / analysis format = nlmny.
;
  define returns      / analysis format = nlmny.
;
  define inventory    / analysis format = nlmny.
;
run ;
```

This selects only a subset of the available columns and sets out the order in which they will appear. A COMPUTED column in PROC REPORT allows us to embed DATA Step syntax within the procedure to augment the existing data on-the-fly, e.g.

```
proc report data = sashelp.shoes nowd ;
  where region = 'Western Europe' ;
  column subsidiary product pctsold sales returns inventory ;
  define subsidiary / order
;
  define product      / display
;
  define pctsold      / computed format = percent8.2  'Percentage/Sold'
;
  define sales        / analysis format = nlmny.
;
  define returns      / analysis format = nlmny.
;
  define inventory    / analysis format = nlmny.
```

Procedures

```
;
compute pctsold ;
    pctsold = (sales.sum - returns.sum) / inventory.sum ;
endcomp ;
run ;
```

The calculation for the COMPUTED column appearing within a COMPUTE / ENDCOMP block referencing fields in the form column.summary-statistic . In this example we have added the COMPUTED column after the product column where we wish it to appear. Submitting this code, however gives a missing value for each row.

The feature which causes this problem is that the calculation in a COMPUTE / ENDCOMP block can only reference columns which have been built as the system reads from left to right, therefore can only include columns to the left of the COMPUTED column. We can demonstrate this by moving the COMPUTED column to the extreme right:

```
proc report data = sashelp.shoes nowd ;
    where region = 'Western Europe' ;
    column subsidiary product sales returns inventory pctsold ;
    define subsidiary / order
;
    define product      / display
;
    define pctsold      / computed format = percent8.2  'Percentage/Sold'
;
    define sales        / analysis format = nlmny.
;
    define returns      / analysis format = nlmny.
;
    define inventory    / analysis format = nlmny.
;
    compute pctsold ;
        pctsold = (sales.sum - returns.sum) / inventory.sum ;
    endcomp ;
run ;
```

The calculations are now performed correctly.

While the calculations may now be correct, if the desire was to position the column immediately after the product column, then we must add in a 'dummy' column and suppress its printing. The following code will give us the desired outcome:

```
proc report data = sashelp.shoes nowd ;
```

Procedures

```
where region = 'Western Europe' ;
column subsidiary product pctsold sales returns inventory _compblock
;
define subsidiary / order
;
define product / display
;
define pctsold / computed format = percent8.2 'Percentage/Sold'
;
define sales / analysis format = nlmny.
;
define returns / analysis format = nlmny.
;
define inventory / analysis format = nlmny.
;
define _compblock / computed noprint
;
compute _compblock ;
    pctsold = (sales.sum - returns.sum) / inventory.sum ;
endcomp ;
run ;
```

Unique solution ID: #1039

Author: Alan D Rudland

Last update: 2017-10-19 11:38