PharmaSUG 2022

Excel the Data

Naga Madeti, Labcorp Drug Development Inc., USA

ABSTRACT

In the pharmaceutical research, when processing the data, it is evident that converting SAS dataset to EXCEL data and EXCEL data to SAS dataset is inevitable. This paper explains coding tips and tricks in using ODS EXCEL with PROC REPORT to generate a beautiful excel sheet, which can be useful to successively import to SAS dataset again if needed without any text wrapping issue for the columns with long text. Also, demonstrates the process of generating EXCEL sheet with addition of multiple tabs. In addition, describes the way for distinguishing the case of not generating the excel sheet due to programming errors from when there is no data.

INTRODUCTION

Carly Fiorina, Ex CEO of Hewlett-Packard, says, "the goal is to turn data into information, and information into insight". An accurate insight of data is possible only when we are able to generate the data into information more precisely. After RTF format of the outputs the most predominantly used output form of SAS datasets is Excel files of the data. There is always a need to generate spread sheet. It may be to send to the data management team to explain the issues of data or to send to a statistician to point out the records where the counts are coming from or to send special reports to find out the impact of certain situation on the study. Sending the data in the excel format allows the reviewers to go through the data with ease and as the excel itself comes with different data analysis methods it makes the process of understanding the data effortless.

SAS DATA TO EXCEL DATA CONVERSION

There are several ways followed to generate Excel sheet from the selected SAS dataset. Some of the methods are as below:

1. Using SAS EXPORT WIZARD
2. Using PROC EXPORT procedure
3. Using the custom made macros - %SAS2XLS.SAS or %SAS2CSV.SAS
4. Using DDE approach in a DATA step
5. Using ODS output with PROC REPORT or PROC PRINT

Among these several approaches, using Output Delivery System (ODS) with PROC REPORT is the flexible approach for generating custom made Excel spread sheet from SAS dataset.

CODE TO EXPORT SAS DATASET WITH ODS OUTPUT USING PROC REPORT

```
data prdsale1;
    length catg $100;
    set sashelp.prdsale;
    where year=1993;
    ntime=floor(actual/predict);
    if predict=0 then catg="No predicetd price";
    if actual<predict then catg="Sold for Less than predicted price";
    if actual=predict then catg="Sold for the predicted price";
    if actual^=predict and ntime=1 then
       catg="Sold for more than the predicted price but less than the double of the predicted price";
    if actual>predict and ntime>=2 then
       catg="Sold for more than or equal the double of the predicted price";

       label catg="Price Category"
             prodtype="Product Type";
run;
```

```
proc freq data=prdsale1;
    tables country*division*prodtype*product*month*catg/out=report(drop=percent);
run;
```

```
proc sort data=report;
    by country division prodtype product catg;
run;
```

```
proc transpose data=report out=report_t(drop=_name_ _label_);
    by country division prodtype product catg;
    id month;
run;
ods excel  file="&_SASWS_./Price Category.xlsx" options(sheet_name="Products_Price" );
```

```
proc report data=report_t ;
    column country division prodtype product catg jan feb mar apr may jun ;
run;
ods excel close;
```

OUTPUT

Label wrapping in to next line

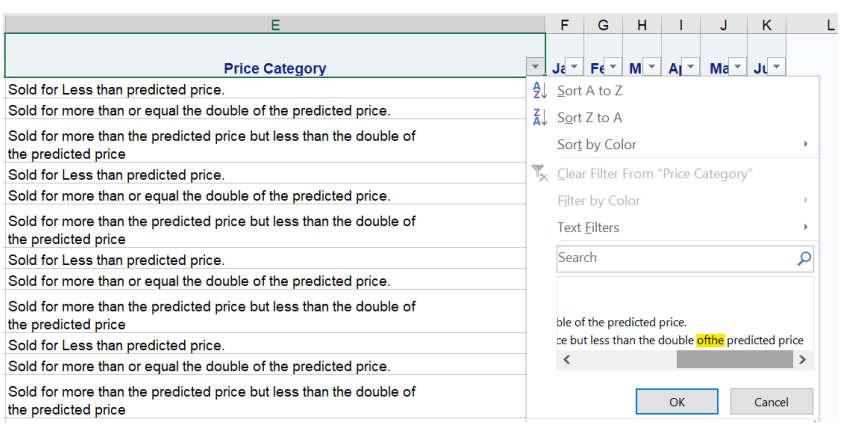
Column text value wrapping in to next line

	A	B	C	D	E	F	G	H	I	J	K	L
1	Country	Division	Product Type	Product	Price Category	Jan	Feb	Mar	Apr	May	Jun	
2	CANADA	CONSUMER	FURNITURE	BED	Sold for Less than predicted price.	2	2	1	.	2	1	
3	CANADA	CONSUMER	FURNITURE	BED	Sold for more than or equal the double of the predicted price.	.	.	.	2	.	.	
4	CANADA	CONSUMER	FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price	.	.	1	.	.	1	
5	CANADA	CONSUMER	FURNITURE	SOFA	Sold for Less than predicted price.	1	2	2	1	1	2	
6	CANADA	CONSUMER	FURNITURE	SOFA	Sold for more than or equal the double of the predicted price.	.	.	.	.	.	.	
7	CANADA	CONSUMER	FURNITURE	SOFA	Sold for more than the predicted price but less than the double of the predicted price	1	.	.	1	1	.	

TEXT WRAPPING ISSUE

A simple code may generate the output, but it is not an accurate way of exporting the data into the excel sheet. Text wrapping is always a real issue at Excel. The issue at Column-E can be observed with in the excel sheet when column values are filtered. we can clearly see how the data is not accurately presented with word wrapping issue by removing the space between the words where it is wrapping in to next line.

In the below screen shot text wrapping issue of removing a space where it wraps in to the next line can be observed by applying filter to the column and selecting the wrapping text.



Also, when the data is imported back to SAS dataset, it is evident that word wrapping caused the changes in text. Sometimes, the data sent to the reviewers might have to be imported again to SAS dataset or to some other format if needed. So, data must be exported without any changes to the text.

```
/**** Importing Excel ****/
 proc import datafile ="&_SASWS_./Price Category.xlsx"
             out= Report_imp
             dbms = xlsx replace;
 run;
```

Product Type	Product	Price Category
FURNITURE	BED	Sold for Less than predicted price.
FURNITURE	BED	Sold for more than the double of the predicted price.
FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price
FURNITURE	SOFA	Sold for Less than predicted price.
FURNITURE	SOFA	Sold for more than the double of the predicted price.
FURNITURE	SOFA	Sold for more than the predicted price but less than the double of the predicted price

Two issues in the imported SAS dataset can be observed predominantly here.

1. The label of the column Product Type is in 2 lines which looks odd.

   Product
   Type

2. At "Price Category" column, it can be observed that the text is wrapping in to next line. By selecting the data where value is in two lines, it can be clearly observed that the space between the words from the last word of the first line and the first word of the second line is removed. This happened due insufficient column width to fit the text when exporting the data.

   KSTRIP(PUT(Price Category, $86.)) = "Sold for more than the predicted price but less than the double ofthe predicted price"

INCREASING THE COLUMN WIDTH TO ELIMINATE WORD WRAPPING ISSUE

PROC REPORT is a best solution to resolve such issues, for it being a more controlling and flexible approach to generate the custom made outputs.

Word wrapping issue can be resolved with STYLE(HEADER) modifier in the define statement for the column where the issue is observed. Using WIDTH= option to increase the column width and setting wrap to no at TAGATTR= can counter the wrapping issue. The width can be modified as per the need with trial and error to best fit the text with in the column without wrapping into the next line.

```
style(header) = { width=15% tagattr="wrap:no"};
```

The code is as below:

```
define prodtype / display 'Product Type'
                  style(header) = { width=15% tagattr="wrap:no"};
```

Similarly, for the long text of Column-E, Price Category-CATG variable, the width of 65% will adjust the column to fit the text in one line avoiding the word wrapping issue. Variable label can also be added with in the define statement as below.

```
define catg / display 'Price Category'
              style(header)={ width=65% tagattr="wrap:no"};
```

The output after these 2 modifications is as below. It can be clearly observed that the text wrapping issue is resolved for the label of Column-C and text values of Column-E.

	A	B	C	D	E	F	G	H	I	J	K	L
1						Year-1993						
2	Country	Division	Product Type	Product	Price Category	Jan	Feb	Mar	Apr	May	Jun	
3	CANADA	CONSUMER	FURNITURE	BED	Sold for Less than predicted price.	2	2	1	.	2	1	
4	CANADA	CONSUMER	FURNITURE	BED	Sold for more than the double of the predicted price.	.	.	.	2	.	.	
5	CANADA	CONSUMER	FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price	.	.	1	.	.	1	
6	CANADA	CONSUMER	FURNITURE	SOFA	Sold for Less than predicted price.	1	2	2	1	1	2	
7	CANADA	CONSUMER	FURNITURE	SOFA	Sold for more than the double of the predicted price.	.	.	.	.	.	.	

ADDITIONAL CHANGES TO SHOW THE COMPATABILITY OF PROC REPORT

Apart from the text wrapping issue, to show the compatibility of PROC REPORT in exporting the data to Excel, additional modifications are addressed to deliver the data in its cleanest format using the flexible PROC REPORT define statements options.

1. Adding group header: Year category for the months will be added as a group header for the months Jan-Jun.

```
column country division prodtype product ord
       ("Year-1993" jan feb mar apr may jun ) ;
```

2. Using format option with in the display statement: The Price Category column values can be formatted to provide extra quality control check of labels to data values instead of IF-THEN-ELSE code. An order variable ORD derived based on the actual price comparison with predicted price and then the data labels are generated as formatted labels with in the PROC REPORT using FORMAT option in the define statement.

```
define ord / display 'Price Category'
             format = $catg.
             style(header) = { width=65% tagattr="wrap:no"};
```

3. The individual cells can be highlighted for the FORMAT of the required condition and using COMPUTE block with BACK GROUND option in PROC REPORT.

```
proc format;
   value hlt
   2 = "lightgreen"
   ;
run;
```

```
compute jan;
call define
(_col_,'style',"STYLE=[BACKGROUND=hlt.]");
endcomp;
```

4. Additional tabs can be added by specifying sheet name, that needs to be added as extra tab, at ODS EXCEL OPTIONS with PROC REPORT procedure above ODS EXCEL CLOSE statement. Any number of additional tabs can be added with in the ODS output statement after PROC REPORT procedure. Additional tabs can be added using PROC PRINT as well in the similar way, as PROC REPORT is compatible to combine with PROC PRINT, PROC SORT, PROC MEANS, PROC FREQ, PROC TABULATE.

CODE TO GENERATE CUSTOMIZED EXCEL SHEET USING PROC REPORT WITH ADDITIONAL TAB

```
proc format;
   value $catg
   "1" = "No predicetd price"
   "2" = "Sold for Less than predicted price"
   "3" = "Sold for the predicted price"
   "4" = "Sold for more than the predicted price but less than the double of the predicted price"
   "5" = "Sold for more than or equal the double of the predicted price"
   ;
   value hlt
   2 = "lightgreen"
   ;
run;
data prdsale1;
    length catg $100;
    set sashelp.prdsale;
    where year=1993;
    ntime=floor(actual/predict);
    if predict=0 then ord ="1";
    if actual<predict then ord="2";
    if actual=predict then ord="3";
    if actual^=predict and ntime=1 then ord="4";
    if actual>predict and ntime>=2 then ord="5";
run;
proc freq data=prdsale1;
    tables country*division*prodtype*product*month*ord/out=report(drop=percent);
run;

proc sort data=report;
    by country division prodtype product ord;
run;

proc transpose data=report out=report_t(drop=_name_ _label_);
    by country division prodtype product ord;
    id month;
run;
proc report data=report_t;
    column country division prodtype product ord
           ("Year-1993" jan feb mar apr may jun ) ;
    define country  / order display;
    define division / order display ;
    define prodtype / display 'Product Type'
                      style(header) = { width=15% tagattr="wrap:no"};
    define product  / order display ;
    define ord / display 'Price Category'
                 format = $catg.
                 style(header) = { width=65% tagattr="wrap:no"};
```

```
    define jan / display ;
    define feb / display ;
    define mar / display ;
    define apr / display ;
    define may / display ;
    define jun / display ;
    compute jan;
    call define
    (_col_,'style',"STYLE=[BACKGROUND=hlt.]");
    endcomp;
run;
ods excel options(sheet_name="Products_Price_2");
 proc report data=report_t;
    column country division prodtype product ord
           ("Year-1993" jul aug sep oct nov dec) ;
    define country  / order display;
    define division / order display ;
    define prodtype / display 'Product Type'
                      style(header) = { width=15% tagattr="wrap:no"};
    define product  / order display ;
    define ord / display 'Price Category'
                 format = $catg.
                 style(header) = { width=65% tagattr="wrap:no"};
    define jul / display ;
    define aug / display ;
    define sep / display ;
    define oct / display ;
    define nov / display ;
    define dec / display ;

run;
```

Cell Highlight

Adding another sheet

OUTPUT

	A	B	C	D	E	F	G	H	I	J	K
1						Year-1993					
2	Country	Division	Product Type	Product	Price Category	Jan	Feb	Mar	Apr	May	Jun
3	CANADA	CONSUMER	FURNITURE	BED	Sold for Less than predicted price	2	2	1	.	2	1
4	CANADA	CONSUMER	FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price	.	.	1	.	.	1
5	CANADA	CONSUMER	FURNITURE	BED	Sold for more than or equal the double of the predicted price	.	.	.	2	.	.
6	CANADA	CONSUMER	FURNITURE	SOFA	Sold for Less than predicted price	1	2	2	1	1	2
7	CANADA	CONSUMER	FURNITURE	SOFA	Sold for more than the predicted price but less than the double of the predicted price	1	.	.	1	1	.
8	CANADA	CONSUMER	FURNITURE	SOFA	Sold for more than or equal the double of the predicted price	.	.	.	.	.	.
9	CANADA	CONSUMER	OFFICE	CHAIR	Sold for Less than predicted price	2	2	1	1	2	2
10	CANADA	CONSUMER	OFFICE	CHAIR	Sold for more than the predicted price but less than the double of the predicted price	.	.	.	.	.	.
11	CANADA	CONSUMER	OFFICE	CHAIR	Sold for more than or equal the double of the predicted price	.	.	1	1	.	.
12	CANADA	CONSUMER	OFFICE	DESK	Sold for Less than predicted price	1	.	1	1	.	1
13	CANADA	CONSUMER	OFFICE	DESK	Sold for more than the predicted price but less than the double of the predicted price	.	2	.	1	.	1
14	CANADA	CONSUMER	OFFICE	DESK	Sold for more than or equal the double of the predicted price	1	.	1	.	2	.
15	CANADA	CONSUMER	OFFICE	TABLE	Sold for Less than predicted price	1	1	1	2	.	1
16	CANADA	CONSUMER	OFFICE	TABLE	Sold for more than the predicted price but less than the double of the predicted price	.	1	.	.	2	.
17	CANADA	CONSUMER	OFFICE	TABLE	Sold for more than or equal the double of the predicted price	1	.	1	.	.	1
18	CANADA	EDUCATION	FURNITURE	BED	Sold for Less than predicted price	2	.	1	1	1	.
19	CANADA	EDUCATION	FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price	.	2	1	1	1	2
20	CANADA	EDUCATION	FURNITURE	BED	Sold for more than or equal the double of the predicted price	.	.	.	.	.	.
21	CANADA	EDUCATION	FURNITURE	SOFA	Sold for Less than predicted price	1	.	1	.	1	.
22	CANADA	EDUCATION	FURNITURE	SOFA	Sold for more than the predicted price but less than the double of the predicted price	1	.	.	1	1	1
23	CANADA	EDUCATION	FURNITURE	SOFA	Sold for more than or equal the double of the predicted price	.	2	1	1	.	1
24	CANADA	EDUCATION	OFFICE	CHAIR	Sold for Less than predicted price	2	.	2	2	1	1
25	CANADA	EDUCATION	OFFICE	CHAIR	Sold for more than the predicted price but less than the double of the predicted price	.	2	.	.	1	.
26	CANADA	EDUCATION	OFFICE	CHAIR	Sold for more than or equal the double of the predicted price						1

Products_Price | Products_Price_2

	A	B	C	D	E	F	G	H	I	J	K
1						Year-1993					
2	Country	Division	Product Type	Product	Price Category	Jul	Aug	Sep	Oct	Nov	Dec
3	CANADA	CONSUMER	FURNITURE	BED	Sold for Less than predicted price	1	1	.	.	1	.
4	CANADA	CONSUMER	FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price	.	1	2	.	1	1
5	CANADA	CONSUMER	FURNITURE	BED	Sold for more than or equal the double of the predicted price	1	.	.	2	.	1
6	CANADA	CONSUMER	FURNITURE	SOFA	Sold for Less than predicted price	1	2	1	.	.	.
7	CANADA	CONSUMER	FURNITURE	SOFA	Sold for more than the predicted price but less than the double of the predicted price	1	.	.	.	1	1
8	CANADA	CONSUMER	FURNITURE	SOFA	Sold for more than or equal the double of the predicted price	.	.	1	2	1	1
9	CANADA	CONSUMER	OFFICE	CHAIR	Sold for Less than predicted price	1	.	2	.	1	1
10	CANADA	CONSUMER	OFFICE	CHAIR	Sold for more than the predicted price but less than the double of the predicted price	.	1	.	1	.	.
11	CANADA	CONSUMER	OFFICE	CHAIR	Sold for more than or equal the double of the predicted price	1	1	.	1	1	1
12	CANADA	CONSUMER	OFFICE	DESK	Sold for Less than predicted price	.	2	2	1	2	1
13	CANADA	CONSUMER	OFFICE	DESK	Sold for more than the predicted price but less than the double of the predicted price	1	.	.	1	.	.
14	CANADA	CONSUMER	OFFICE	DESK	Sold for more than or equal the double of the predicted price	1	.	.	.	.	1
15	CANADA	CONSUMER	OFFICE	TABLE	Sold for Less than predicted price	1	2	2	2	2	1
16	CANADA	CONSUMER	OFFICE	TABLE	Sold for more than the predicted price but less than the double of the predicted price	1	.	.	.	.	1
17	CANADA	CONSUMER	OFFICE	TABLE	Sold for more than or equal the double of the predicted price	.	.	.	.	.	.
18	CANADA	EDUCATION	FURNITURE	BED	Sold for Less than predicted price	2	2	1	1	1	1
19	CANADA	EDUCATION	FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price	.	.	.	.	1	1
20	CANADA	EDUCATION	FURNITURE	BED	Sold for more than or equal the double of the predicted price	.	.	1	1	.	.
21	CANADA	EDUCATION	FURNITURE	SOFA	Sold for Less than predicted price	1	2	.	1	2	1
22	CANADA	EDUCATION	FURNITURE	SOFA	Sold for more than the predicted price but less than the double of the predicted price	.	.	1	1	.	1
23	CANADA	EDUCATION	FURNITURE	SOFA	Sold for more than or equal the double of the predicted price	1	.	1	.	.	.
24	CANADA	EDUCATION	OFFICE	CHAIR	Sold for Less than predicted price	1	1	.	.	.	1
25	CANADA	EDUCATION	OFFICE	CHAIR	Sold for more than the predicted price but less than the double of the predicted price	1	1	.	.	1	1
26	CANADA	EDUCATION	OFFICE	CHAIR	Sold for more than or equal the double of the predicted price	.	.	2	2	1	.
27	CANADA	EDUCATION	OFFICE	DESK	Sold for Less than predicted price	1	.	1	1	1	.
28	CANADA	EDUCATION	OFFICE	DESK	Sold for more than the predicted price but less than the double of the predicted price	1	1	.	1	1	.
29	CANADA	EDUCATION	OFFICE	DESK	Sold for more than or equal the double of the predicted price	.	1	1	.	.	2
30	CANADA	EDUCATION	OFFICE	TABLE	Sold for Less than predicted price	.	1	1	1	2	1
31	CANADA	EDUCATION	OFFICE	TABLE	Sold for more than the predicted price but less than the double of the predicted price	.	.	1	.	.	.
32	CANADA	EDUCATION	OFFICE	TABLE	Sold for more than or equal the double of the predicted price	2	1	.	1	.	1
33	GERMANY	CONSUMER	FURNITURE	BED	Sold for Less than predicted price	1	1	1	1	1	1
34	GERMANY	CONSUMER	FURNITURE	BED	Sold for more than the predicted price but less than the double of the predicted price	.	.	1	.	1	.

Products_Price | Products_Price_2

WHEN NO DATA…

When there is no data in the required selection of dataset, excel will just display a blank page or it will not generate a sheet which is the same result when there is some issue due to programming error. These two scenarios can be distinguished using a no data condition. At the REPORT_T dataset, a sheet will be generated if there are any products sold for exactly the predicted price, which is category with ORD="3". Based on the number of records the output can be generated as a normal PROC REPORT output or no records sheet will be generated with a selected text for the no records as below.

```
data report_eq;
    set report_t;
    where ord="3";
run;
```

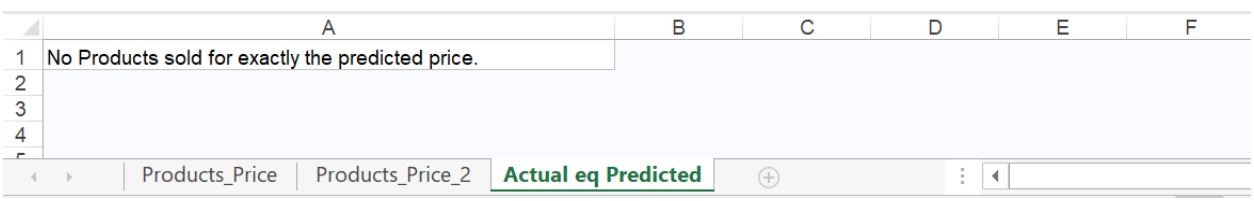
```
proc sql noprint;
   select count(*) into:norec from report_eq;
quit;

%put &norec;
```

The sheet "Actual eq Predicted" is added as an additional tab if there are any records and if not then the no records text is displayed.

```
ods excel options(sheet_name="Actual eq Predicted");

%if &norec>0 %then %do;
 proc report data=report_eq;
    column country division prodtype product ord
           ("Year-1993" jan feb mar apr may jun ) ;
    define country  / order display;
    define division / order display ;
    define prodtype / display 'Product Type'
                      style(header) = { width=15% tagattr="wrap:no"};
    define product  / order display ;
    define ord / display 'Price Category'
                 format = $catg.
                 style(header) = { width=65% tagattr="wrap:no"};
run;
%end;

 %else %do;
 data report_eq;
    col = "No Products sold for exactly the predicted price.";
 run;

 proc report data=report_eq  noheader;
 run;
 %end;

 ods excel close;
```

As there are no products that are with actual price equal to predicted price, the text "No Products sold for exactly predicted price." is displayed at the excel output instead of a blank sheet.

OUTPUT



CONCLUSION

To generate excel output from SAS dataset, using ODS output with PROC REPORT is the flexible and effective way. The common issue of text wrapping while generating excel sheet from SAS data including long text columns can be resolved with PROC REPORT options. A more beautiful and customized excel output can be obtained using PROC REPORT's diverse functionality.

ACKNOWLEDGMENTS

Author would like to thank the employer Labcorp Drug Development Inc. for the approval to present this paper at PharmaSUG.

REFERENCES

1. Jane Eslinger. "The SAS Programmers PROC REPORT Handbook", SAS Institute Inc. 2016.
2. David-Erick Lafontant. "Proc report- a better way to display data". Available at https://uisug.org.uiowa.edu/sites/uisug.org.uiowa.edu/files/slide3.pdf
3. Eric Gebhart. "ODS ExcelXP: Tag Attr Is It! Using and Understanding the TAGATTR= Style Attribute with the ExcelXP Tagset". Proceedings of the SAS Global Forum 2010. Available at https://support.sas.com/resources/papers/proceedings10/031-2010.pdf

CONTACT INFORMATION:

Naga Madeti, Labcorp Drug Development Inc.

naga.madeti@labcorp.com