SAS® to Excel: ODS Tagsets to Hundred Files or Hundred Tabs

Abdus Shahid, Western Oregon University

ABSTRACT

This paper will present methods of outputting dozens or hundreds of tables to an excel workbook as tabs or to hundreds excel files with the use of ODS Tagsets that comes with a table of contents. All the formatting for the tables are done with SAS codes, so that little or no manual formatting is needed once the final output is opened in excel. The method described is especially useful when a large number of tables are needed to be complied in one excel workbook.

Some knowledge of Base SAS coding and XML output is required to fully utilize this process.

INTRODUCTION

SAS® outputs do not always look visually appealing and sometimes need to be used in conjunction with other software to produce formatted tables and reports. ODS Tagsets in SAS® provides a convenient method for this which interfaces well with excel. This paper shows procedure to do that and details that can make an excel file with many tabs useful for reports.

RETENTION AND GRADUATION

Retention (and/or persistence) and graduation are two key metrics used to evaluate performance of an institution of higher education. Retention is a measure of the rate at which students who are new to the institution return for subsequent terms. Persistence is defined as a measure of the rate at which all students return for subsequent terms whether or not they are new to the institutions. This paper will present examples of retention and persistence of first-time full-time cohorts only.

Western Oregon University institutional research office has data on cohorts from 2004. These data sets are merged with enrollment data from 2005 to 2018 to find out enrollment status of cohorts which provides retention and persistence data in subsequent years. However, institutions are also interested in retention and persistence by groups such as gender, race/ethnicity, financial aid groups, resident or non-resident groups, and sub-groups among these groups such as retention of female Hispanic students. So, the office of Institutional Research and Effectiveness produced retention and persistence tables for 51 categories.

With cohort, enrollment, group and sub-group identifications, data to be used for these tables look like the following:

	Cohort	Enrolled_2004	Enrolled_2005	Enrolled_2006	Enrolled_2007	Sex_Code	Resident_code	Resident_Sex_Code
1	2004	1	.	.	.	1	1	1
2	2004	1	.	.	.	2	1	2
3	2004	1	1	.	.	2	1	2
4	2004	1	1	1	.	1	1	1
5	2004	1	.	.	.	1	1	1
6	2005	.	1	1	1	1	2	3
7	2005	.	1	.	.	2	2	4
8	2005	.	1	.	.	2	2	4
9	2005	.	1	1	1	2	2	4
10	2005	.	1	.	.	1	2	3

Figure 1. Student Data

To look at retention and persistence of all the groups and sub-groups, first a loop is defined for all 51 tables in the following manner:

```
/************************ Start loop ********************************/
%let title1=All;              %let var1=all_code;      %let value1=1;

%let title2=Male;             %let var2=sex_code;      %let value2=1;

%let title3=Female;           %let var3=sex_code;      %let value3=2;

%let title5=Resident;         %let var5=resident_code; %let value5=1;

%let title6=Non_resident;     %let var6=resident_code; %let value6=2;

%let title7=URM;              %let var7=urm_code;      %let value7=1;

%let title8=First_Generation; %let var8=first_gen_code; %let value8=1;

%let title9=Pell_Eligible;    %let var9=pell_code;     %let value9=1;

%let title10=Legacy;          %let var10=Legacy_code; %let value10=1;

%let title44=Legacy_Male;%let var44=legacy_sex_code;  %let value44=1;

%let title45=Legacy_Female;%let var45=legacy_sex_code;%let value45=2;
.
.
.
.
%let title46=Resident_Male;   %let var46=resident_sex_code; %let value46=1;

%let title47=Resident_Female; %let var47=resident_sex_code; %let value47=2;

%let title50=Rural_School;    %let var50=rural_code;       %let value50=1;

%let title51=Non_Rural_School; %let var51=rural_code;       %let value51=2;

%let beginLoop=1;

%let endLoop=51;

/*************************** End loop ********************************/
```

Next, PROC TABULATE procedure is used to tabulate retention and persistence:

```
%Macro Tabulate;

%do i=&beginloop %to &Endloop;
ods tagsets.excelxp options(sheet_interval='none' sheet_name="&&Title&i.."
embedded_titles='yes' frozen_headers='Yes' Skip_spaces='1,0,0,0,1'
absolute_column_width='19,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4'
center_horizontal='yes' scale='80' embedded_footnotes='Yes' dpi='650');

proc tabulate data=retention_2 format=comma12. out = out_&&Title&i..;
class ir_cohort;
var Enrolled_2004 - Enrolled_2018;
table ir_cohort='', (enrolled_2004='2004' enrolled_2005='2005'
enrolled_2006='2006'
                                enrolled_2007='2007' enrolled_2008='2008'
enrolled_2009='2009' enrolled_2010='2010'
                                enrolled_2011='2011' enrolled_2012='2012'
enrolled_2013='2013' enrolled_2014='2014'
                                enrolled_2015='2015' enrolled_2016='2016'
enrolled_2017='2017' enrolled_2018='2018')
                                /printmiss box={label="Cohort/Year"};
keylabel sum=' ';
where FT = 1 and &&var&i=&&value&i;
run;
Title;

*percent of cohort;
data percent_&&Title&i..;

set out_&&Title&i..;

%do j = 2004 %to 2018;

if ir_cohort = &j. then cohort = enrolled_&j._sum;

%do k = 1 %to 1;

percent_%eval(&j+&k.) =((enrolled_%eval(&j.+&k.)_sum)/cohort)*100;

%end;

%end;

run;

%end;

%Mend Tabulate;

%Tabulate
```

Notice that out files, highlighted in red from the PROC TABULATE process are also being produced. Retention tables need to show not only the headcounts but also the percentages of headcount of the cohort by year. The out file will be used to manipulate the output from PROC TABULATE process for that purpose.

After manipulating the data in a few steps and formatting for final output, I run the following codes:

```sas
%let output =I:\Shahid\SCARF\Output\Cohort\Retention and Graduation;

%Macro Print;

ods tagsets.excelxp file="&output\Retention and Persistence V13 E.xml"
    options(absolute_column_width='14,7,7,7,7,7,7,7,7,7,7,7,7,7,7,7,7,7,7,7'
                    Orientation            ='landscape'
                    FitToPage              ='yes'
                    Pages_FitWidth         ='8'
                    Pages_FitHeight        ='80'
                    Embedded_Titles        ='yes'
                    Embedded_Footnotes     ='yes'
                    Frozen_Headers         ='1'
                    Column_Repeat          ='2-4'
/*                  Row_Repeat             ='1-4'*/
/*                  Autofit_Height         ='yes'*/
/*                  Sheet_name             ='Owner ID Match'*/
            row_repeat='header'
                    Sheet_interval= 'print'
                    Contents               ='yes'
                    Index                  ='yes'
                    Print_Footer = "&LIR, Western Oregon University &CPage &P
of &N
               &R&D ")
            style=shahid4;


%do i=&beginloop %to &Endloop;

proc print data = Final_A_&&Title&i.. noobs;
title "&&Title&i..";

ODS TAGSETS.EXCELXP
     OPTIONS (sheet_name="&&Title&i..");
     var Cohort Fall_2004 - Fall_2018/ style(header)={just=r}
style(data)={just=r};
ods proclabel "&&Title&i..";
footnote1 'Source: SCARF and Saturn Research tables';
footnote2 j=c color=blue height=10pt link="#Contents!A1" "Return to
Contents";

run;

%end;

run;

title;
footnote1;
footnote2;

ods tagsets.excelxp close;

%Mend Print;

%print
```

OPTIONS

Many options are available to customize output tables according to one's needs. For example, absolute column width will determine width of all the columns if only one length is defined but this option can be used to define all the columns of the tables. Table of contents is produced by using the option

Contents ='yes', but can be turned off by 'no'. Orientation = 'landscape' can be changed to 'portrait'.

One important option is Sheet_interval. Setting Sheet_interval= 'print' will send all the tables in separate tabs. If Sheet_interval is set to 'none', then all the tables will be complied in one tab only. There are several key words available for this option that sends tables to separate tabs, like print, bygroup, table, page and proc. The default is none.

All the tables need not to have the same format. Formats for tables can be changed in the following way:

```
ODS TAGSETS.EXCELXP
      OPTIONS (Sheet_Name = "Class"
                          absolute_column_width='8');
proc print data = class;
title1 'Class Data';
run;


ODS TAGSETS.EXCELXP
      OPTIONS (Sheet_Name = "Test 2"
                          absolute_column_width='10,8,6,6,10');
proc print data = test;
title1 'Test Data';
title2 'Year 2018';
run;
```

ODS Tagsets allow us to customize data presented in the tables by variables. Here is an example:

```
ODS TAGSETS.EXCELXP
      OPTIONS (Sheet_Name = "Cohort 2018");

ods proclabel "First-time Full-time Cohort: 2018";
footnote1 'Source: WOU_EDW_COHORT_DATA';
footnote2 j=c color=blue height=10pt link="#Contents!A1" "Return to
Contents";
*end;

*cohort 2017;
proc print data = cohort.cohort_2017 noobs;
title 'First-time Full-time Cohort: 2017';
var _NAME_ Total;
var Percent /style(column)={tagattr="format:##0.0%"};
```

Tables also need not be in the same file. Assigning individual file names will produce separate file for each table.

OUTPUT

Tagsets produce .xml files which can be saved in excel.

It is possible to output multiple files directly to excel:

```
%let outputloc = I:\Shahid\8_AD_HOCS\Output\Tina_Max;

%macro export;

%do i = 2017 %to 2015 %by -1;

proc export
  data=wou_&i.
  dbms=xlsx
  outfile="&outputloc.\Cohorts 2017 - 2015.xlsx"
  replace;
  sheet = "&i.";
run;

%end;

%mend export;

%export
```

However, options available for .xml files are not available in this case.

CONCLUSION

ODS Tagsets is a great way to present outputs from SAS® in a formatted manner in one file which can be made convenient to navigate. The options available for this process makes a repetitive work easy.

REFERENCES

- *Base SAS® Procedures Guide*
- An important resource on ODS Excel XP options can be found here:

  http://support.sas.com/rnd/base/ods/odsmarkup/TipSheet_ods_xl_xp.pdf

CONTACT INFORMATION

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

Abdus Shahid
Western Oregon University
503-838-9331
shahida@wou.edu
Web: http://www.wou.edu/institutionalresearch/

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.