

# Illustrating generation of MSWord tables via HTML with the %cattable and %contable macros

Matthew T. Karafa, PhD., Cleveland Clinic Foundation, Cleveland, Ohio

## ABSTRACT

The author presents two macros for use in SAS version 7.0 or higher. The %cattable macro takes a list of any number of categorical factors, and %conttable takes a similar listing of continuous factors. Both macros present the association of the list with a 2-level response variable and provide summary statistics taken from several procs with many options for table appearance and statistic selection including odds ratios and appropriate significance levels from both parametric and non-parametric methods. The author uses these two macros, which generate manuscript-ready descriptive summary tables, to demonstrate how to take Microsoft Word® mock-up tables and generate SAS code that produces customizable tables from ODS output datasets and data step code. Used in conjunction with Microsoft Word®'s ability to read and interpret HTML files these allow additional flexibility in formatting and layout often difficult to achieve using the ODS RTF destination.

## Introduction

Often the goal is to have SAS generate a specific table with a very specified look, for example in use in a multi-center trial annual report. However, customizing the ODS RTF template to achieve specific results can be cumbersome, tables with complex designs with spanning columns and rows for example. Since html is a plain-text markup language, SAS can generate html code that contains MSWord specific tags to create a highly customizable table.

The current pair of macros illustrate the standard "Table 1" that is typically seen in a given manuscript. From a fairly simple macro call they can create tables with odds ratios, appropriate summary statistics, and p-values. In addition, both create MSWord compliant html, which can be easily converted to a standard word document.

## Categorical Measures - %cattable()

The first of the two macros takes a response variable, `GroupVar`, and compares two of its groups across one or more categorical variables given in `vlist`. If `GroupVar` is not binary, the macro will compare the minimum value to the maximum value with the minimum used as the reference group. In order to calculate the odds ratios, a reference group must be determined. By default the macro uses the last group alphabetically based on the formatted value of the group. Thus one can force the reference group simply by numbering the formats 1,2,3,... Z.

For example consider this data set called Test:

ID	Group	Male	Black	Handedness	Small	Age	Cost	LOS	Satis
1	0	0	0	2	0	43.1	973	7.7	3
2	0	0	0	2	0	72.3	1011	3.5	3
...									
144	1	0	0	2	0	59.5	1031	23.1	1
145	1	0	0	1	0	68.1	985	443.5	2

With the following sample call :

```
%CatTable(ds=Test, GroupVar = Group, vlist =Male Black Handedness Small,
  DefaultStyle=D, DefaultCutoff=0.15, ColP=0, OddsRatios= 1,
  ListingFile=T, DelimitedFile=F, HtmlFile=T,
  Ofile_Html=./MWSug2006.html, owrite_html = T
  MainTitle=%str(Table 1: Comparison of Group 1 to Group 0),
);
run;
```

the macro generates some output to our list file or output window, and the file MWSug2006.html. The macro will take the data from TEST and use all previously defined labels and formats for the listed dataset. MWSug2006.html is overwritten each time the macro is run. Notice that the

MainTitle parameter needs to be surrounded by the macro function %str() to preserve spaces. Similarly the ofile\_html would need %str() were spaces a part of the file name. Opening MWSug2006.html in MSWord, Table 1 looks like:

Table 1: Comparison of Group 1 vs. Group 0

Factor	Level	Total	Risk Group		Reference Group		Odds Ratio (CI)	P value
			N	(%)	N	(%)		
Male							9.3 (3.6 ,24.2)	<0.001
	Z. Female	70	6	8.6	64	91.4		
	1. Male	75	35	46.7	40	53.3		
Black							1.02 (0.30 ,3.4)	0.99F
	Z. White	131	37	28.2	94	71.8		
	1. Black	14	4	28.6	10	71.4		
Handedness								0.86F
	1:Left	8	3	37.5	5	62.5	1.6 (0.32 ,7.5)	
	Z:Right	43	12	27.9	31	72.1	1.0 (REF)	
	2:Both	94	26	27.7	68	72.3	0.99 (0.44 ,2.2)	
Small							N/A	<0.001F
	Z. Non-Event	132	28	21.2	104	78.8		
	1. Event	13	13	100.0	0	0.00		

Where the “F” after the P-value denotes the Fisher’s Exact Test.

The listing of available parameters are:

Required Parameters (parameters required or usually input by a user).

- GroupVar The grouping variable, typically a binary variable, but non-binary factors will compare minimum to maximum.
  - Vlist The list of variables to be analyzed via standard Chi-Square or an Exact Test. You may specify many variables separated by spaces. Each may have some, all or none of the switches set. (For switch descriptions see below)
- For each variable you may specify a "/" and a style for that factor. If you specify nothing, the DefaultStyle will be used. You can specify:

- /X : Forces the Pearson Chi-square test
- /E : Gives Fisher’s Exact Test
- /D : Determine’s Chi-square or Fisher’s based on some cutoff (Default = 25%)
- /cut: Percent of Expected cells with count less than 5 to use to Fisher’s Exact

So for example to force the variable “sex” to be Exact one specifies “sex/E” in the vlist. To force it to be Chi-Square it would be sex/X. To force it to be determined by 10% of cells with expected counts below 5 it would be “sex/D/5”

Optional Parameters.

- ds The data set to be analyzed. (Default is \_LAST\_)
- SampleWt Define the variable containing sampling weight as in the weight statement. (Default = -1, no weight)
- MainTitle Text to be used for the Main title.
- DefaultStyle Style to use in if no Style is given to a specific variable. In this case it sets as either X: Chi Square E: Fishers Exact Test, or D: Determine (Default is D, determine.)
- DefaultCutoff Determines the cut off percent of cells with expected counts below 5 before using Fishers Exact Test. (Default is 25.)
- ColP Specifies whether you want COLUMN percents. Expects 0,1. (Default is 0 - Row Percents)
- OddsRatios Toggles the Odds Ratios in the table (Default is 0, No OR)
- MonteCarloLimit Sets the default number of categories before the exact method uses the Monte Carlo Estimate of the Fishers Exact Test. (Default is 3 categories)
- ListingFile Toggles report generation to the SAS list file/output window (Default=T)
- DelimitedFile Toggles report generation to the delimited file (Default=T)
- Owrite Specifies whether to overwrite the delimited file. Expects T,F,Y or N. (Default is F no overwrite)
- Ofile Name of the file you would like the delimited text to go to. Quotes optional. (Default is "./Temp.txt")

SepChar	Character you would like to use as a delimiter in your delimited file. Must be in quotes or hex notation for special characters. (Default is '09'x (Tab))
HtmlFile	Toggles report generation to the Html file (Default=T)
owrite_html	Specifies whether or not to overwrite the Html file. Expects T,F,Y or N. (Default is F no overwrite)
ofile_html	Name of the file you would like the Html text to go to. Quotes optional. (Default is "/Temp.html")
DEBUG	Toggles the full output (e.g. turns ods listing close off) and macro error debugging routines. Expects 1,0 input (Default is 0, Debug off.)

### Continuous Measures - %conttable()

In a similar manner, %conttable() takes a response variable, GroupVar, and compares two of its groups across one or more continuous variables given in vlist. This macro can create either mean (sd) with a T-test comparison, or median (25<sup>th</sup>, 75<sup>th</sup> percentiles) with a Wilcoxon test comparison. GroupVar is handled exactly like in %cattable(). However with continuous measures we have an additional complication in the odds ratio, the unit increase. For most continuous measures the odds ratio is expressed in terms of a unit increase of the listed factor, however it is common to express the odds ratios as reflecting a larger unit change. For this we can specify a unit by placing a "+" and a number after the variable name like "age+10" and "cost+25" in the example below.

Consider again the data in Test, with this sample call:

```
%conttable(ds=Test, GroupVar = Group, vlist = age+10 los cost/w+25 satis/w,
  DefaultStyle = T, DefaultTtype = U, OddsRatios = 1,
  qtldecplace = 1, Estdecplace = 1, Unitdecplace = 1,
  ListingFile = T, DelimitedFile = T, HtmlFile = T,
  ListingFile=T, DelimitedFile=F, HtmlFile=T,
  Ofile_Html=../MWSug2006.html, owrite_html = T
  MainTitle=%str(Table 2: Comparison of Group 1 to Group 0 [cont.]),
);
```

this Table would be the result:

Table 2: Comparison of Group 1 vs. Group 0 [cont.]

Factor	Risk Group		Reference Group		Units	Odds Ratio (CI)	P value
	N	Statistics	N	Statistics			
Age	41	37.7(20.3)	104	36.4(21.8)	10	1.03 (0.87 ,1.2)	0.72S
LOS	41	136.7(230.0)	104	168.7(258.7)	1	1.00 (1.00 ,1.00)	0.47S
Cost	41	992.0 (975.0, 1024.0)	104	1001.0 (972.5, 1017.0)	25	1.07 (0.81 ,1.4)	0.79W
Satis	41	3.0 (3.0, 3.0)	104	3.0 (3.0, 3.0)	1	1.01 (0.72 ,1.4)	0.97W

%Conttable()'s available parameters are:

Required parameters for the %conttable() macro.

GroupVar	The grouping variable, non-binary factors will compare minimum to maximum.
Vlist	The list of variables to be analyzed via t-test or the Wilcoxon test. You may specify many variables separated by spaces. Each may have some, all or none of the switches set. (For switch descriptions see below)

For each variable you may specify a "/" and a style for your var. If you specify nothing, the Default Style will be used. You can specify:

/W	: Gives median (q1,q3) for each variable and the Wilcoxon Test
/T	: Gives t-test. You can further specify /U, /D, or /E
/U	: Unequal Variances
/E	: Equal Variances
/D	: Determine the variance based on the F-test. Cutoff alpha default is P<0.10
+Number	: Gives the Units for the given variable for OR calculations. Must be the last switch.

So for example Specifying:

age : use DefaultTest, and if T-test is specified as default the unequal variances (Satterwaith)  
age/W : gives the Wilcoxon test and median (25<sup>th</sup>, 75<sup>th</sup> percentiles)  
age/T : gives the T-Test, and mean(sd) using the Default T-type  
age/T/U : gives the T-Test, and mean(sd) and forces the Unequal variance  
age/T/E : gives the T-Test, and mean(sd) and forces the Equal variance  
age/T/D : gives the T-Test, and mean(sd) and determines the variances using the F-test P<0.10  
age/T/D/15: gives the T-Test, and median(sd) and determines the variances using the F-test P<0.15  
age+10 : Specifies increase of 10 for odds ratio calculations  
age/W+10 : gives the Wilcoxon test and median (q1,q3) and specifies increase of 10 for odds ratio calculations

Optional parameters for the %conttable() macro.

ds The data set to be analyzed. (Default is \_LAST\_)  
DefaultStyle Style to use in if no Style is given to a specific variable. Expects W (Wilcoxon) or T (t-test) (Default is T for t-test)  
DefaultTtype Assuming you are using the t-test this determines the type of t-test (E: Equal, U: Unequal, or D: Determine) (Default is U)  
PrintTvalue Prints the value of the T statistic (Default is 0, No)  
OddsRatios Outputs Odds Ratios (Default is 0, No) Default the 'Units' for a given variable is 1.0  
Maintitle Text to be used for the Main title.  
Qtldeplace Number of decimal places for the 25<sup>th</sup>, 75<sup>th</sup> percentiles. Expects 0-9. (Default = 1)  
Estdeplace Number of decimal places for mean (SD) and/or median. Expects 0-9. (Default = 1)  
Unitdeplace Number of decimal places for odds ratios "units". Expects 0-9. (Default = 1)  
overwrite Specifies whether to over write the output file. Expects Y or N. (Default is N no overwrite)  
ofile Name of the file you'd like the delimited text to go to. Must be in quotes. (Default is "./Temp.txt")  
SepChar Character you'd like to use as a delimiter in your delimited file. Must be in quotes or hex notation. (Default is '09'x (Tab))  
HtmlFile Toggles report generation to the Html file (Default=T)  
overwrite\_html Specifies whether or not to overwrite the Html file. Expects T,F,Y or N. (Default is F no overwrite)  
ofile\_html Name of the file you wouldd like the Html text to go to. Quotes optional. (Default is "./Temp.html")  
DEBUG Toggles the full output (e.g. turns ods listing close off) and macro error debugging routines. Expects 1,0 input (Default is 0, Debug off.)

### Determining the HTML to be created

In the macro code for each of these two macros, there is a subroutine that generates the HTML code. The basic methodology is to use the data step to create a file and "put" our HTML code into the file by brute force. Constructing such a data step is a fairly simple task in SAS the general structure:

```
data _null_;  
  file "SomeFileName.Ext";  
  put "Line to be put to the output file";  
  ...  
  put "This line is last";  
run;
```

However, the question becomes what tags and text should be put out to make the HTML table. To determine the needed HTML code, the first step is to create a mock up of the table we wish to create in MSWord. This can be as simple or as complex as we like so long as the cells of interest can easily be determined. In the mock table, generic terms that can later be replaced with SAS macro variables are used rather than actual values. Next, save this file using "Save as Web Page" and open the resulting HTML file in your favorite text editor.

For example %conttable will generate a table similar to the one below. In this mockup we replace the actual values with placeholders for the data we would want placed in the cells. For the current discussion, the text to be replaced is highlighted in red:

<u>MainTitle</u>							
	<u>Gp1_Label</u>		<u>Gp2_Label</u>				
Factor	N	Statistics	N	Statistics	Units	OR (CI)	P value
<u>var_label_i</u>	<u>Gp1SampSz_i</u>	<u>Gp1Stats_i</u>	<u>Gp2SampSz_i</u>	<u>Gp2Stats_i</u>	<u>Units_i</u>	<u>OR_i</u>	<u>Pvalue_i</u>

FootNote

A search through the resulting HTML document finds a line that has the form:

```
<p class=MsoNormal>
<span style='font-size:10.0pt;mso-bidi-font-size:12.0pt'>
_MainTitle_
</span>
</p>
```

so in the sas data \_null\_ we would need the following command:

```
put "<p class=MsoNormal>"
put "<span style='font-size:10.0pt;mso-bidi-font-size:12.0pt'>"
put _MainTitle_
put "</span></p>";
```

If we store the title to be used in a dataset variable called “\_Maintitle\_”, the data \_null\_ step will replace that with its value in the resulting text file. Note also that the tags can be split up for legibility. Thus the reader of the program can see the HTML tags that are put into the textfile as well as the SAS dataset variables that are used in the process. Searching for the other placeholders and constructing a similar series of put statements can generate the remainder of the table.

## Conclusion

Like most programming tasks, there are several paths that can be taken to achieve a goal. Though much of the functionality of this method is available if a user were to specify a custom template, using HTML can offer a much more WYSIWYG method for generating tables that have to be done frequently. With some effort on the front end, a production level table can be generated directly from SAS meeting very complex details that can be read in both MSWord or directly on the web.

## References

SAS Institute Inc. (1999), SAS OnlineDoc® documentation, Version 8, Cary NC

SAS Institute Inc. (2002), SAS OnlineDoc® documentation, Version 9, Cary NC

For copies of the macros go to <http://www.geocities.com/mkarafa/Stats/Macros.html>

## Contact Information

Matthew T. Karafa, PhD  
 Quantitative Health Sciences, Cleveland Clinic Foundation  
 9500 Euclid Avenue  
 Cleveland, Ohio 44195  
 Phone: (216) 445 – 9556  
 Fax: (216) 444 - 8021  
 Email: [karafam@ccf.org](mailto:karafam@ccf.org)

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.