DOCUMENTATION FOR: Table_Macros_DrexelBSC.sas

AUTHOR: Janeway Granche, jg3463@drexel.edu
LAST UPDATED: 2020 Mar 5
DEVELOPED FOR: SAS 9.4
LICENSE: MIT License – see License.txt
SEE ALSO: Table_Macros_DrexelBSC_Example_1.1.2.sas for examples of macros use

DOCUMENT OVERVIEW:

1. SAS file contents
2. Macros descriptions
3. Notes
4. Methods for use
5. Macro details
6. Example output

SAS FILE CONTENTS:

MACROS:

1. SetTableParameters -> Sets global table parameters
2. TableNs -> Count or count (%) data for table heading
3. Count_Pct -> Count (%), p-value from Chi Sq test
4. Mean_SD -> Mean (SD), p-value from ANOVA F-test
5. Median_IQR -> Median [Q1-Q3], p-value from Kruskal-Wallis Chi Sq test

SUB-MACROS:

6. Set_HasFmt -> Determines existence of format, based the FindFmt macro available at: http://support.sas.com/kb/24/019.html
7. Round_Counts -> If weights used, round counts to whole numbers
8. Set_Count_Pct -> combines count and percent data
9. Set_RowName -> Sets output table RowName values
10. Set_RowOrder -> Sets output table RowOrder values
11. Round_Pval -> Rounds Pval and removes leading 0s

MACROS DESCRIPTIONS:

SetTableParameters sets all table-wide parameters and must be run first. Macro checks that certain parameter values are valid and outputs table settings to log. Only TableData and VarHeadings are required. Omitting ByBar results in a single-column table. PctType can be set to COL (default), ROW, PCTONLY, or NONE to get column %, row %, only column percent data, or only count data. P-values are omitted by default but can be included by setting

PvalDigits > 0. Macros will round p-values to PvalDigits places (max=4) and remove leading zeros. Weights are omitted by default. Statistical tests are not implemented for weighted data.

TableNs creates a single row dataset with count or count (%) data with columns for each category in ByVar (if using) plus the Total.

Count_Pct creates a dataset with one row for each category of the variable Var containing count and/or (%) data for the total, plus each category of ByVar (if using). Also output are the variable total VarN and columns containing data that will determine how table entries are combined and displayed. For dichotomous variables, set parameter RowDisplay equal to the value of the variable to be displayed for single row output (not compatible with PctType=ROW). If PvalDigits > 0, macro will calculate the Var*ByVar Chi Square p-value (default) or the Fisher's Exact Test p-value if Fisher is set to 1. Statistical tests are not implemented for weighted data. All percents are rounded to nearest whole number unless using PctOnly setting, in which they are rounded to 1 decimal place.

Mean_SD creates single row datasets containing mean (SD) data for each category in ByVar (if using), in addition to table formatting data, and variable total VarN. If PvalDigits > 0, Mean_SD will include the Var*ByVar ANOVA F-Test p-value. If weights are specified, Mean_SD standard deviations are calculated using Proc Means with the VarDef=WDF option. Statistical tests are not implemented for weighted data.

Median_IQR creates single row datasets containing median [Q1-Q3] data for each category in ByVar (if using), in addition to table formatting data, and variable total VarN. If PvalDigits > 0, output will include the Kruskal-Wallis Chi Square p-value. If weights are specified, medians are computed with Proc Means using the QMethods=OS option. Statistical tests are not implemented for weighted data.

NOTES:
Global Variables: _TableData, _VarHeadings, _PctType, _ByVar, _ByVarFmt, _PvalDigits, _Weight, _TableSet, _HasFmt. _HasFmt is used in several macros and is deleted at the end of

each use. The rest are table parameters and are set and validated in SetTableParameters macro.

Temporary Datasets: In addition to the output dataset, each macro creates several temporary datasets in the Work library, which are deleted before the macro terminates. The all temporary dataset names start with an underscore, so users can avoid having their datasets overwritten by refraining from starting dataset names with underscores.

Missing Values: Missing values are automatically omitted from macro tables. In order to include a Missing category, I suggest using a version of the variable with missing values reset to "Missing" or "99", etc.

Default Formats: If categorical variable (in)formats are omitted, macros will first look for a format and informat with the same name as the corresponding variable, and if they exist, will use them to set RowName, RowOrder, and the column names. If no RowOrderFmt is provided for a character variable, the macro will print a warning, and RowOrder will be set to 1 for all variable categories.

METHOD FOR USE:
1. Create required format: the variable headings format converts order in which the variables are to appear in the final table (set by TabOrder in each variable macro call) to the variable heading names to be displayed. For the table Ns heading, the input should be set to 0.
2. Create informats and formats: Informats can be used to set variable category display order and are required for character variables. Formats can set variable and ByVar category display names and are recommended for numeric variables.
3. Run SetTableParameters setting table data, variable headings format, data display format, and other table options.
4. Run TableNs to get table heading Ns.
5. Run Count_Pct on each categorical variable and Mean_SD or Median_IQR on each continuous variable to be included in table.
6. Merge macro output data by TabOrder and RowOrder.

7. Print combined table using Proc Report. Users can apply additional formatting in Proc Report, or they can output the assembled table data to a spreadsheet for further processing.

See Table_Macros_DrexelBSC_Example.sas for examples of macro use and table production. See also annotated example output at the end of this document.

MACRO DETAILS:

1. SetTableParameters

MACRO CALL

%SetTableParameters(TableData =,
            VarHeadings =,
            ByVar =,
            ByVarFmt =,
            PctType =,
            PvalDigits =,
            Weight=
)

REQUIRED PARAMETERS:

- TableData = dataset used in table
- VarHeadings = format converting table order to variable headings

OPTIONAL PARAMETERS:

- ByVar = grouping variable that defines column. Omit for single column tables
- ByVarFmt = format that determines how ByVar category names are displayed. If not set, macro uses format with same name as ByVar, if it exists
- PctType = categorical output format. Possible values: Col, Row, PctOnly, and None. Default (Col) sets table data to count (col%) format
- PvalDigits = p-values can be rounded to 1 – 4 digits or can be omitted entirely (default)
- Weight = Var weights, omit for unweighted data

OUTPUT: Table settings are output to SAS log.

2. TableNs

MACRO CALL

    %TableNs(Output=,

            TabOrder=

)

REQUIRED PARAMETERS: NONE

OPTIONAL PARAMETERS:

- Output – set to Counts, Pct, or CntPct for count, percent, or count (percent) output. Default: Counts
- TabOrder – sets order in which output is tup

OUTPUT: a dataset DemoTab[TabOrder] containing the following variables:

- RowName - set by VarHeadings format
- Total variable, plus one variable for each category in ByVar (if set), names determined by ByVarFmt format (if set), containing count or count (%) data, depending on Pct value
- TabOrder - an integer indicating the order in which the variable Var will appear in the final table, also used by VarHeadings format. Default value = 0

3. Count_Pct

MACRO CALL

    %Count_Pct(Var=,

            TabOrder=,

            RowDisplay =,

            RowOrderFmt=,

            RowNameFmt=,

            Fisher=

)

REQUIRED PARAMETERS:

- Var - a categorical variable (numeric or character) that will appear in the table
- TabOrder - an integer indicating the order in which the variable Var will appear in the final table, also used by VarHeadings format

OPTIONAL PARAMETERS:

- RowOrderFmt – (Required for character data) an informat that converts Var categories into integers defining the order in which characters will be displayed. Numbering should start at 1 since 0 is reserved for variable headings. If parameter is not set in macro, the default value is an informat with the same name as Var. If such informat does not exist, numerical variables will be ordered as stored, while character variable categories are output with RowOrder = 1, and macro will output warning message.

- RowDisplay - value of Var category to be displayed for single-row output. Omit to display all Var categories.

- RowNameFmt - a format that converts Var categories to categorical headings to be displayed on table. If omitted, macro will look for a format with same name as Var and use it if it exists. Otherwise, categories are displayed as are stored.

- Fisher – Set to 1 to get Fisher's Exact Test p instead of Chi Squared p (default). Only applicable if PvalDigits > 0.

OUTPUT: a dataset DemoTab[TabOrder] containing the following variables:

- RowName - Var category names set by RowNameFmt format and the table section heading set by VarHeadings format

- variable Var containing levels of the input Var, which should match up to names in RowName

- one variable for each category in ByVar (if using), names determined by ByVarFmt format (if using), plus a Total variable, containing count and/or (%) data, format set by PctType

- TabOrder - integer set by macro parameter TabOrder determining order in which variables appear in final table.

- RowOrder - integers set by RowOrderFmt informat determining order in which Var categories will appear in final table

- Pval - the Chi Squared or Fisher's Exact test p-value, heading row only, if PvalDigits > 0

- VarN - total number of observations with valid Var and ByVar entries, heading row only

4. Mean_SD

MACRO CALL

```
%Mean_SD(Var=,
         TabOrder=,
```

RoundingUnit=
)

REQUIRED PARAMETERS:

- Var - a numeric variable that will appear in the table
- TabOrder - an integer indicating the order in which the variable Var will appear in the final table, also used by VarHeadings format

OPTIONAL PARAMETERS:

- RoundingUnit - the level of precision to which means will be rounded using the SAS ROUND function. Default value: .1

OUTPUT: single-row dataset DemoTab[TabOrder] containing the following variables:

- RowName - the variable heading name set by the VarHeadings format from the TabOrder value
- Total, plus one variable for each category in ByVar (if using), names determined by ByVarFmt format (if using), containing Mean (SD) data
- TabOrder - constant integer determining order in which variables appear in final table
- RowOrder - set to 0
- Pval - the ANOVA F-test p-value, included if PvalDigits > 0
- VarN - total number of observations with valid Var and ByVar entries

5. Median_IQR

MACRO CALL

%Median_IQR(Var=,
           TabOrder=,
           RoundingUnit=
)

REQUIRED PARAMETERS:

- Var - a numeric variable that will appear in the table
- TabOrder - an integer indicating the order in which the variable Var will appear in the final table, also used by VarHeadings format

OPTIONAL PARAMETERS:

- RoundingUnit - the level of precision to which medians and quartiles 1 and 3 will be rounded using the SAS ROUND function. Default value: .1

OUTPUT: single-row dataset DemoTab[TabOrder] containing the following variables:

- RowName - the variable heading name set by the VarHeadings format from the TabOrder value

- Total, plus one variable for each category in ByVar (if using), names determined by ByVarFmt format (if using), containing Median [Q1-Q3] data

- TabOrder - constant integer determining order in which variables appear in final table

- RowOrder - set to 0

- Pval - the Kruskal-Wallis Test p-value from PROC NPAR1WAY, included if PvalDigits > 0

EXAMPLE OUTPUT

Annotated Table1 data from Proc Print, line 89.

%SetTableParameters, Line 49, sets ByVar (sets columns), VarHeadings (sets variable headings), PvalDigits (set pval column), and table dataset

Set in each macro call

Set by Chol_order

Original variable

Obs	Alive	Dead	Total	RowName	tabOrder	rowOrder	Chol_Status	VarN	pval
1	3218 (62)	1991 (38)	5209 (100)	Participants, n (%)	0	0		.	
2				Cholesterol	1	0		5057	<.001
3	998 (32)	407 (21)	1405 (28)	Desirable	1	1	Desirable	.	
4	1186 (38)	675 (35)	1861 (37)	Borderline	1	2	Borderline	.	
5	951 (30)	840 (44)	1791 (35)	High	1	3	High	.	
6	951 (30)	840 (44)	1791 (35)	High Cholesterol	2	0	High	5057	<.001
7				Cholesterol, categories renamed	3	0		5057	<.001
8	998 (32)	407 (21)	1405 (28)	Under XXX	3	1	Desirable	.	
9	1186 (38)	675 (35)	1861 (37)	Between XXX and YYY	3	2	Borderline	.	
10	951 (30)	840 (44)	1791 (35)	Over YYY	3	3	High	.	
11				Cholesterol, unordered	4	0		5057	<.001
12	1186 (38)	675 (35)	1861 (37)	Borderline	4	1	Borderline	.	
13	998 (32)	407 (21)	1405 (28)	Desirable	4	1	Desirable	.	
14	951 (30)	840 (44)	1791 (35)	High	4	1	High	.	
15	221.96 (43.07)	236.32 (46.47)	227.42 (44.94)	Cholesterol, mean (SD)	5	0		5057	<.001
16	217 [192-246]	232 [204-264]	223 [196-255]	Cholesterol, median [IQR]	6	0		5057	<.001

%TableNs, Line 60

%Count_Pct, Line 63

%Count_Pct, Line 67 with RowDisplay=High
%Count_Pct, Line 70

%Count_Pct, Line 73

%Mean_SD, Line 76

%Median_IQR, Line 79

Count (Col%) data displayed by default

Headings defined in Tab1VarHead, Line 23, using tabOrder numbers set in each macro call

Names set by RowNameFmt = Chol_Name

Omitting RowOrderFmt results in all 1s

Final Table 1, produced by Proc Report, line 93. Export to spreadsheet for further formatting.

Table 1. Exploration of Macro Functionality				
	Alive	**Deceased**	**Total**	**p***
Participants, n (%)	3218 (62)	1991 (38)	5209 (100)	
Cholesterol				<.001
Desirable	998 (32)	407 (21)	1405 (28)	
Borderline	1186 (38)	675 (35)	1861 (37)	
High	951 (30)	840 (44)	1791 (35)	
High Cholesterol	951 (30)	840 (44)	1791 (35)	<.001
Cholesterol, categories renamed				<.001
Under XXX	998 (32)	407 (21)	1405 (28)	
Between XXX and YYY	1186 (38)	675 (35)	1861 (37)	
Over YYY	951 (30)	840 (44)	1791 (35)	
Cholesterol, unordered				<.001
Borderline	1186 (38)	675 (35)	1861 (37)	
Desirable	998 (32)	407 (21)	1405 (28)	
High	951 (30)	840 (44)	1791 (35)	
Cholesterol, mean (SD)	221.96 (43.07)	236.32 (46.47)	227.42 (44.94)	<.001
Cholesterol, median [IQR]	217 [192-246]	232 [204-264]	223 [196-255]	<.001
Note: * Chi Square test for count data unless otherwise noted, ANOVA F test for Mean (SD) data, Kruskal-Wallis Chi Square test for Median [IQR] data.				

Final Table 2, fully formatted in Proc Report syntax lines 216 – 271.

Table 2 - Characteristics for Deceased Subjects by Cause of Death

Characteristics	Cancer	Cerebral Vascular Disease	Coronary Heart Disease	Other	Unknown	Full Sample[a]
Number of participants	539	378	605	357	112	1991
Age at entry, mean (SD)	47.3 (8)	50.8 (7.1)	48.8 (7.9)	48.6 (8.6)	50 (8.6)	48.8 (8.1)
Age at death, mean (SD)	68.9 (10.2)	71.9 (10.4)	70.3 (9.7)	71.3 (12.1)	72.6 (10.8)	70.5 (10.6)
Sex (female), %	47.5	52.6	34.5	49.3	50	45
Height (in), mean (SD)	65.2 (3.3)	64.4 (3.5)	65.5 (3.6)	64.6 (3.7)	64.6 (3.6)	65 (3.6)
Weight (lbs), mean (SD)	155 (27)	158 (31)	164 (30)	155 (31)	154 (33)	158 (30)
Smoking status, %						
Non-smoker	43.7	48.5	45.2	44.8	42.7	45.2
Light (1-5)	8.6	10.5	9.2	9.9	10.9	9.5
Moderate (6-15)	10.8	11.5	10.2	9.9	14.5	10.8
Heavy (16-25)	23.5	19.6	21.2	25.8	23.6	22.5
Very Heavy (> 25)	13.4	9.9	14.2	9.6	8.2	12
Blood pressure, %						
Optimal	13.2	7.4	5.3	9.2	8	8.7
Normal	39.3	25.9	29.4	34.2	32.1	32.4
High	47.5	66.7	65.3	56.6	59.8	58.9
Diastolic BP, mean (SD)	85.7 (11.9)	92.1 (16.9)	91.7 (13.8)	89.1 (14.4)	89.3 (14)	89.5 (14.3)
Systolic BP, mean (SD)	138 (23)	154 (34)	149 (26)	144 (28)	146 (24)	146 (28)
Cholesterol, %						
Desirable	25.3	22	14.6	26.2	17.6	21.2
Borderline	38.8	33.9	33.6	33.7	34.3	35.1
High	35.9	44.1	51.8	40.1	48.1	43.7
Cholesterol, mean (SD)	228 (43)	235 (45)	247 (50)	231 (43)	240 (48)	236 (46)

[a] Variables with missing data have the following totals: height, 1986; weight, 1988; smoking status, 1971; cholesterol, 1922.

Final Table 3, produced by Proc Report, line 374. Export to spreadsheet for further formatting.

Table 3. Ice Cream Study Counts with Weighted Percents						
	$10 and More		Under $10		Total Sample	
Sample Characteristics	**n**	**%**	**n**	**%**	**N**	**%**
Number of Participants	17		23		40	
Population Represented by Sample	1387	43.9	1776	56.1	3163	100
Grade 9	5	29.1	6	27.2	11	28
Grades						
7	3	16.4	17	72.8	20	48.1
8	9	54.5			9	23.9
9	5	29.1	6	27.2	11	28
Spending, Mean (SD) (weighted)	14 (3.1)		5 (2.7)		8.9 (5.3)	
Spending, Med [IQR] (weighted)	14 [11-16]		6 [2-7]		8 [4-13]	
Note: percents represent weighted values.						