Descriptive Statistics Summary Table

Macro Name: DESCRIPTIVE
Created Date/Author: Feb. 2012/Yuan Liu
Last Update Date/Person: Oct 31, 2018/Yuan Liu
Other Contributors: Dana Nickleach
Current Version: V16
Working Environment: SAS 9.4 English version

Contact: Dr. Yuan Liu Yliu31@emory.edu

Purpose: To produce a descriptive statistics summary table for each variable in the dataset. The frequency, including # of missing value will be generated for categorical variables; and summary statistics (n, mean, median, Q1, Q3, min, max, standard deviation, # of missing) for numerical variables.

Notes: 1) When using this macro to generate DATA DICTIONARY or DATA CODEBOOK, set *DICTIONARY=T*. The data dictionary contains variable name, label, formatted and unformatted levels for categorical variables, and descriptive statistics for all variables. 2) The order of variables shown in the final summary report is the original order in the database, and it will be better prepare covariates into the desired order in the data management step (see http://support.sas.com/kb/8/395.html). 3) A copy of SAS dataset of the final summary report can be saved by the name specified in OUTDATA. Users can use the information in this dataset to generate the report in a customized format.

Citation:
Yuan Liu, Dana Nickleach, Chao Zhang, Jeffrey Switchenko, Jeanne Kowalski. *Carry Out Streamlined Routine Data Analyses with Reports for Observational Studies: Introduction to a Series of Generic SAS ® Macros.* Oct 2018. Preprint. DOI:10.13140/RG.2.2.25265.84329

Parameters:

Macro variable	Description	Required
DATASET	The name of the data set to be analyzed.	Yes
CLIST	List of categorical variables, separated by empty space.	Yes
NLIST	List of numerical variables, separated by empty space.	Yes
CHI	Set to T to calculate the chi-square goodness of fit p-value for categorical variables. The default value is F.	No
NONPAR	Set to T to calculate the sign test p-value for median=0 for numeric variables (optional). The default value is F.	No
TTEST	Set to T to calculate the t-test p-value for mean=0 for numeric variables, (optional). The default value is F.	No
DICTIONARY	Set to T to generate a Data Dictionary or Data Codebook of dataset.	No
DOC	Set to T to create a RTF file containing the output or F to suppress creation of the RTF file (optional). The default value is T. Set to F will be useful to see a preliminary results in SAS without opening Word.	No
OUTDATA	If not empty, output the final report as a SAS dataset by the name specified.	No
OUTPATH	Path for output table to be stored.	Yes

ORIENTATION	Value of PORTRAIT or LANDSCAPE to indicate the page layout of the report. The default value is PORTRAIT.	No
FNAME	File name for output table.	Yes
FOOTNOTE	Text of footnote to include in the table (optional). It should be in quotes. Leave this field blank if not including a footnote.	No
DEBUG	Set to T if running in debug mode (optional). Work datasets will not be deleted in debug mode. This is useful if you are editing the code or want to further manipulate the resulting data sets. The default value is F.	No

Usage Example:

```
DATA analysis;
   input id os_censor Sex $ Age duration os progress $ trt $;
   LABEL os = 'Overall Survival (months)'
      progress = 'Progression'
      trt = 'Treatment'
      duration = 'Duration of Radiation';
   DATALINES;
1    1    M    40    44    20    No    B
2    1    F    45    46    16    Yes   A
3    1    F    40    32    20    No    B
4    1    F    47    32    23    No    B
5    0    M    41    25    22    No    B
6    1    M    54    35    13    No    B
7    1    M    48    50    9     Yes   A
8    1    M    36    33    12    Yes   B
9    0    F    49    51    8     Yes   A
10   1    M    49    52    10    Yes   A
11   1    M    44    35    12    No    A
12   1    M    49    50    8     Yes   A
13   1    M    44    44    14    Yes   A
14   1    M    50    31    10    Yes   A
15   1    M    53    40    15    No    B
16   0    M    52    29    20    Yes   B
17   1    F    46    45    5     Yes   A
18   1    F    37    44    11    Yes   A
19   1    M    49    46    13    No    B
20   1    M    42    31    11    No    A
;

PROC format; value trt 1 = "A" 2= "B";run;
DATA analysis;set analysis;format trt trt.;run;

TITLE 'Table 1 Descriptive Statistics';
%DESCRIPTIVE(dataset = analysis,
      clist = trt sex,
      nlist = age duration,
      outpath = C:\Documents and Settings\User\My Documents\,
      fname = Table 1 Descriptive Statistics);
TITLE;

TITLE 'Table 2 Data Dictionary';
%DESCRIPTIVE(dataset = analysis,
      clist = trt sex,
      nlist = age duration,
      dictionary = T,
      outpath = C:\Documents and Settings\User\My Documents\,
      fname = Table 2 Data Dictionary);
TITLE;
```

Summary Table Example:

Table 1 Descriptive Statistics

Variable	Level	N (%) = 20
Treatment	A	11 (55.0)
	B	9 (45.0)
Sex	F	6 (30.0)
	M	14 (70.0)
Age	Mean	45.75
	Median	46.50
	Minimum	36.00
	Maximum	54.00
	Std Dev	5.17
	Missing	0.00
Duration of Radiation	Mean	39.75
	Median	42.00
	Minimum	25.00
	Maximum	52.00
	Std Dev	8.41
	Missing	0.00

Table 2 Data Dictionary

Variable Name	Variable Label	Level		N (%) = 20
Sex	Sex	F	F	6 (30.0)
		M	M	14 (70.0)
Age	Age		Mean	45.75
			Median	46.50
			Minimum	36.00
			Maximum	54.00
			Std Dev	5.17
			Missing	0.00
duration	Duration of Radiation		Mean	39.75
			Median	42.00
			Minimum	25.00
			Maximum	52.00
			Std Dev	8.41
			Missing	0.00
trt	Treatment	1	A	11 (55.0)
		2	B	9 (45.0)

Permission:

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell

Log of Updates:

Date	By	Description	Version
6/11/ 12	Dana Nickleach (dnickle@emory.edu)	Changed PROC DATASETS so that the file being analyzed will not be deleted.	V3
6/19/ 12	Dana Nickleach (dnickle@emory.edu)	Got rid of leading space in RTF file name.	V3
7/30/ 12	Dana Nickleach (dnickle@emory.edu)	Prevented deletion of temporary data sets not created by the macro, added debug parameter, and restored options at the end of the macro to those set at the call of macro	V4
8/9/12	Dana Nickleach (dnickle@emory.edu)	Increased length of variable to 256, the maximum possible length of a variable label so that labels are not cut off in the table and added the SPANROWS option to accommodate longer labels.	V5
8/31/12	Dana Nickleach (dnickle@emory.edu)	Made an adjustment so that a continuous and categorical variable with the same label can be displayed in the table.	V6
10/22/12	Dana Nickleach (dnickle@emory.edu)	Prevented ODS listing error, prevented case sensitivity for T/F parameters, and prefixed data set names with underscores to avoid interfering with user data sets.	V7
11/12/12	Dana Nickleach (dnickle@emory.edu)	Fixed bug preventing character values of "-" from printing in the table.	V8
2/25/13	Dana Nickleach (dnickle@emory.edu)	Fixed error when only numeric variables are supplied, made adjustments so that the order of the variables is preserved even when categorical and numerical variables have the same label, added CHI and DOC parameters, and prevented printing of preliminary steps.	V9
5/10/13	Dana Nickleach (dnickle@emory.edu)	Added FOOTNOTE parameter.	V10
1/2/14	Dana Nickleach (dnickle@emory.edu)	Changed the length of the level variable to 200, the maximum length that is output with ODS, so that values of level are not truncated.	V11

2/6/14	Dana Nickleach (dnickle@emory.edu)	Fixed with in level column to prevent the table from wrapping to the next page.	V12
2/20/14	Dana Nickleach (dnickle@emory.edu)	Modified so that the chi-square test is only calculated if requested.	V13
7/9/14	Dana Nickleach (dnickle@emory.edu)	Added TTEST and NONPAR parameter.	V14
6/25/15	Yuan Liu (yliu31@emory.edu)	Added DICTIONARY parameter; merge N and % column into one;	V15
10/31/2018	Yuan Liu (yliu31@emory.edu)	Added ORITENTATION, OUTDATA parameter; the order of covariation in the final report will be the same as in the dataset; for numerical variable added Q1 and Q3.	V16