**DOCUMENTATION: table1 Macro**

**Macro Name:** TABLE1

**Created Date/Author:**January 2019/Debra Goldman ([dgstat](https://github.com/DGStat/) on GitHub)

**Last Update Date/Person**: Oct 1, 2021/Hannah Kalvin & Stephanie Lobaugh

**Other Contributors:** Hannah Kalvin & Stephanie Lobaugh

**Current Version (reflected in tag on github):** table1\_v1.0

**Working Environment:** SAS 9.4

# Contact

Hannah Kalvin [kalvinh@mskcc.org](mailto:kalvinh@mskcc.org)

Stephanie Lobaugh [lobaughs@mskcc.org](mailto:lobaughs@mskcc.org)

# 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. This macro creates descriptive statistics one variable at a time and requires a separate macro call for each variable.

# Notes:

In order to use this macro, variables must be numeric; to print items that are character, create format that can be applied in table1 macro call. Nonparametric testing is the default. If you would like to not call each variable individually, see table1\_loop macro documentation.

# Parameters:

|  |  |  |
| --- | --- | --- |
| Macro variable | Description | Required |
| DATA | The name of the data set to be analyzed. | Yes |
| ROWVARTYPE | 1: categorical stats; 2: continuous summary stats | Yes |
| ROWVARFORMAT | format of variable. both numeric and character formats will work. note: can also use formats such as 5.0 for count data or $ formats for categorical data  numeric can be set to percentn, integer, decimal etc. | Yes |
| CREATEDATA | set to 1 in **first call of table1** macro to create dataset to save information to (not necessary if you already have a shell dataset created outside of macro call)  default is 0 | No but recommended |
| TABLE\_NAME | provide table name from create table1 macro above. If not provided, defaults to table1\_combined | No but recommended |
| GROUPVAR | grouping variable. If blank, it runs single column. If variable listed, it will provide an overall and values stratified by this grouping variable. default is to provide columnar proportion (see below) | No but recommended |
| GROUPVARFORMAT | grouping variable format. Will apply this format to variable. If group variable is not provided,this variable is not needed. | No but recommended |
| GROUPPERCENT | percentage type included in combined variable  all percents will be available in the dataset  1=column percent; 2=row percent; 0=overall percent  default is 1, column percent | No |
| TEST | test for statistical significance  0 = do not test; 1 = test  default is 1, test for significance | No |
| INCLUDE\_MISS | 1: include missing values; 0: exclude missing values.  default is 1 (include missing)  if missing are excluded, also excludes the following  for character variables: ("-999","999","-888","888","missing","unknown","n/a","na");  for numeric variables: (-999,999,-888,888) | No |
| CONTCOUNT | 1: include N for continuous var in category label for non-grouped data or as a secondary row for grouped data. label is not an option for grouped data ; 0: exclude print  note that the frequency will still be included in the output dataset  2: include the N as a secondary row with N= as the variable value  default is 1 (include N with Median(range) / Median (IQR)) | No |
| CONTSTAT | 1 for median and range in combined var; 2 for median and IQR in combined variable  3 for mean and SD  default is 1 (use median and range) or anything other than 0 will produce the range  note that all stats are included in output dataset. this simply refers to combined presentation | No |
| ORDER | order of categorical variable output  freq for descending frequency; internal for unformatted; external for formatted order  data for data order  default is freq (order by descending frequency) | No |
| COMBOFORMAT | format of combined variable. applies to categorical variables  1: percent to 1 decimal and 2 spaces between freq and pct;  2: percent to integer with % sign with 2 spaces between freq and pct;  3: percent to 1 decimal with % sign with 2 spaces betweens freq and pct;  4: percent to 1 decimal with one space between freq and pct  default is 1 ( N (PCT.1)) | No |
| VIEW | output of table  1: output table; 0 or any other value, no table  default is blank | No |
| NAMEFORMAT | give variable format name for dataset; default is blank (no format) | No |
| DELETEDAT | for troubleshooting, whether or not you want to delete or keep working datasets (default=1)  0= keep working datasets; 1= delete working datasets | No |

# Usage Example:

FILENAME utility URL "https://raw.githubusercontent.com/slobaugh/create\_msk\_SAS\_project/main/utility.sas";

%INCLUDE utility;

\*read in data;

libname test "G:\Macros\Testing\";

**data** d1;

set test.sample\_data;

**run**;

**proc** **contents** data=d1;

**run**;

**proc** **print** data=d1 (obs=**20**);

**run**;

\*create formats for data;

**proc** **format**;

value marf **1**="Yes"

**0**="No";

value hivposf **0**="No"

**1**="Yes";

**run**;

\*use table1 macro;

%***table1***(createdata=**1**,data=d1,rowvar=married,rowvartype=**1**,rowvarformat=marf.);

%***table1***(createdata=**0**,data=d1,rowvar=hivpos\_n,rowvartype=**1**,rowvarformat=hivposf.);