

Finally a Proc Tabulate I can use to QC any Table
Sunil Gupta, SASSavvy.com

		Total
Age	N	19
	Mean	13.3
	Std	1.5
	Median	13.0
	Min	11.0
	Max	16.0
	Q1	12.0
	Q3	15.0
Sex		
F	N	9
	%	47.4
M	N	10
	%	52.6
Subtotal	N	19
Height	N	19
	Mean	62.3
	Std	5.1
	Median	62.8
	Min	51.3
	Max	72.0
	Q1	57.5
	Q3	66.5
Weight	N	19
	Mean	100.0

* Ultimate Proc Tabulate syntax to qc tables with categorical and continuous variables;

```
proc print data=sashelp.class; run;
```

* Close ODS LISTING to create tables in RTF;
ods listing close;

* Set wide and long enough settings;
options linesize=256 pagesize=5000;

* Apply style to improve table appearance;
ods rtf file='C:\qc\v_t_demo.rtf';
title 'QC Demo';

* Group continuous variables into categories;
* Apply MULTILABEL option if overlap of values, use with MLF CLASS option;
* Apply PICTURE to display % with percentages;
* Apply NOTSORTED for user control order instead of default alpha sort order;

```
*proc format;  
* value agegrp 0 - 50 = '0 to 50'  
          51 - 100 = '51 to 100';
```

```
* picture mypct (round) low-high='009%';  
*quit;
```

* Presort the dataset by categorical variables to control display order of columns and rows;
* Apply WHERE condition to subset dataset as needed;

```
proc sort data=sashelp.class out=class;  
  by sex;  
run;
```

* ORDER=DATA or ORDER=FORMATTED to display columns/rows based on data or formatted labels with PROC FORMAT;
* MISSING to include all missing values else the denominator is not correct;
* FORMAT= for global format of statistics;
* Apply combination of options to display or not display missing values - MISSING, PRINTMISS, MISSTEXT and PRELOADFMT;

```
proc tabulate data=class order=data missing format=5.1;
```

```
* List of all continuous variables;  
var age height weight;
```

```
* Apply BY statement for by page tables;  
* by sex;
```

```
* Apply formats as needed;  
* format age agegrp.;  
* format age mypct.;
```

* List all categorical variables with the PRELOADFMT and PRINTMISS option to display all possible values even if missing in data;

```
* MLF option with MULTILABEL Proc Format option if values overlap;  
* Can have separate CLASS statements for different variables for different options;  
class sex/preloadfmt;
```

- * Assign all statistics labels;
keylabel N='N' PCTN='% ' COLPCTN='% ' MEAN='Mean' MEDIAN='Median' MIN='Min' MAX='Max' STD='Std' Q1='Q1'
Q3='Q3' ALL='Subtotal';
- * PRINTMISS with PRELOADFMT option to display all values;
- * MISSTEXT to display 0 instead of '.' for zero counts;
- * Note that for categorical variables with missing values, the percentages may be incorrect because the denominator include missing counts;
- * ALL for grand total else grouping variable to create columns;
- * TABLE statement has a combination of continuous and categorical variables in the order of the table to qc;
- * Use variable labels as default but can override, ex. age='Age (yrs)';
- * Add more row variables as needed as new lines;
- * For continuous variables, keep standard list and order of statistics - n mean std median min max q1 q3;
- * For continuous variables, can select other stats also - nmiss;
- * For categorical variables, keep n and COLPCTN to calculate column percentages, ALL is added as subtotals - sex*(n colpctn) all;
- * For categorical variables, can select any percent direction calculation - COLPCTN, ROWPCTN, REPPCTN or PAGEPCTN;
- * For nesting, cross two or more variables in row or column dimension;
- * For indenting, apply nesting and INDENT=4 option;
- * Valid TABLE statement assures number of columns is the same for each variable syntax;

```

tables (
  age*(n*f=3. mean std median min max q1 q3)
  sex*(n*f=3. colpctn) all
  height*(n*f=3. mean std median min max q1 q3)
  weight*(n*f=3. mean std median min max q1 q3)
), (all='Total')/ rts=75 printmiss misstext='0';
run;
ods rtf close;

```

- * Correct percentages due to missing values;
- * Create dummy variable to track non-missing and missing;

```

data class
set class;

```

```

if sex="" then sex_cnt=0; else sex_cnt=1;
run;

```

```

proc tabulate data=class order=data missing format=5.1;

```

```

class sex/preloadfmt;
var sex_cnt age height weight;

```

```

keylabel N='N' PCTN='% ' COLPCTN='% ' MEAN='Mean' MEDIAN='Median' MIN='Min' MAX='Max' STD='Std' Q1='Q1'  
Q3='Q3' ALL='Subtotal';

```

```

tables (
  sex

```

```
), sex_cnt*(n colpctsum)/ rts=75 printmiss misstext='0';
```

```
* For variable with unique values that create multiple columns;
```

```
* tables (
```

```
  sex
```

```
), (&colvar*sex_cnt)*(n colpctsum)/ rts=75 printmiss misstext='0';
```

```
run;
```