Objective: to create a summary table Age by Sex for 'Mean(SD)' and 'Median(Min, Max)’

/\* --- First, the summary results were extracted to a SAS dataset: --\*/

**proc** **means** data=sashelp.class n mean std median min max maxdec=**1**;

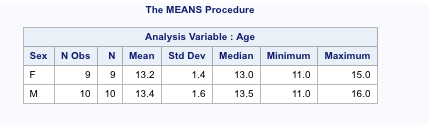
class Sex;

var Age;

output out= \_out

n=nAge mean=mAge std=sdAge median=medAge min=minAge max=maxAge;

**run**;



**proc** **print** data=\_out;

**run**;



/\* Then the mean and median variables were created by concatenating: \*/

**data** \_out2;

set \_out;

meanSD = strip(round(mAge,**0.1**))||" ("||strip(round(sdAge,**0.1**))||")";

medianMM = strip(round(medAge,**0.1**))||" ("||strip(round(minAge,**0.1**))||

", "||strip(round(maxAge,**0.1**))||")";

**run**;

**proc** **print** data=\_out2;

**run**; 

/\*------------------------------------------------------------------\*/

**Constructing the table:**

Method 1: Created a dataset for each variable (wide-to-long) and then combined

/\* Created one dataset for the first variable "Variable": \*/

**data** want1a(keep=Variable);

/\* created an array to hold the three values for variable 'Variable': \*/

array ar2{**3**} $**20** \_temporary\_ ("Age, year"," Mean (SD)"," Median (Min, Max)");

do i=**1** to **3**; /\* create three obs \*/

Variable = ar2{i};

output;

end;

**run**;

**proc** **print** data=want1a;

**run**;



/\* Created one dataset for the variable "Overall": \*/

**data** want1o(keep=Overall);

set \_out2 (where=(\_type\_=**0**));

array ar{**2**} $ meanSD medianMM;

length Overall $**18**;

Overall = ""; /\* create the first obs \*/

output;

do i=**1** to **2**; /\* create the rest two obs \*/

Overall = ar{i};

output;

end;

**run**;

**proc** **print** data=want1o;

**run**;



/\* Created one dataset for the variable "Female": \*/

**data** want1f(keep=Female);

set \_out2 (where=(Sex="F"));

array ar{**2**} $ meanSD medianMM;

length Female $**18**;

Female = ""; /\* create the first obs \*/

output;

do i=**1** to **2**; /\* create the rest two obs \*/

Female = ar{i};

output;

end;

**run**;

**proc** **print** data=want1f;

**run**;



/\* Created one dataset for the variable "Male": \*/

**data** want1m(keep=Male);

set \_out2 (where=(Sex="M"));

array ar{**2**} $ meanSD medianMM;

length Male $**18**;

Male = ""; /\* create the first obs \*/

output;

do i=**1** to **2**; /\* create the rest two obs \*/

Male = ar{i};

output;

end;

**run**;

**proc** **print** data=want1m;

**run**;



/\* now combine all four datasets horizontally: \*/

**data** want1;

set want1a;

set want1o;

set want1f;

set want1m;

**run**;

**proc** **report** data=want1 style(column)={asis=on};

column Variable Overall ('Sex' Female Male); /\* spanning header (or super-header?) \*/

**run**;



Method 2: use array/do-loops for Long-to-wide transpose in one data step

**data** want2(keep=Variable Overall Female Male);

length Variable $**20**; /\* this must be long enought to hold " Median (Min, Max)" \*/

/\* Created these three new variables in this array: \*/

array sTable{**3**} $**16** Overall Female Male;

/\* The order of these three variables in this array must be in the same order

as the observations in variable "Sex" of dataset \_out2!! \*/

/\* -- want to create three observation, so each do-loop need an 'output' --\*/

/\* The first observation: \*/

do i = **1** to **3**;

/\* creates the first observation to hold value 'Age, year' for variable 'Variable': \*/

Variable = "Age, year";

/\* then set other three variables to missing for this obs: \*/

set \_out2;

sTable{i} = "";

end;

output; /\* need 'output' and only 'ouput' after the 'do-end' loop \*/

/\* creates the second observation to hold value ' Mean (SD)' for variable 'Variable' and

Mean (SD) for other three variables: \*/

do i = **1** to **3**;

Variable = " Mean (SD)";

set \_out2;

/\* other three variable will take the value from variable 'meanSD': \*/

sTable{i} = meanSD;

end;

output;

/\* creates the 3rd observation to hold value ' Median (Min, Max)' for variable

'Variable' and Median (Min, Max) for other three variables: \*/

do i = **1** to **3**;

Variable = " Median (Min, Max)";

set \_out2;

/\* other three variable will take the value from variable 'medianMM': \*/

sTable{i} = medianMM;

end;

output;

**run**;

**proc** **print** data=want2 noobs style(data)={asis=on};

**run**;

Method 3: use Proc Transpose:

/\* re-assign the values for 'Sex' before transpose it: \*/

**data** want3a;

length Sex $**7**;

set \_out2;

/\* re-assign the values for this variable: \*/

if Sex ='' then Sex = "Overall";

else if Sex = "F" then Sex = "Female";

else if Sex = "M" then Sex = "Male";

**run**;

/\* Now transpose it to long form: \*/

**proc** **transpose** data=want3a out=want3b name=V0;

var meanSD medianMM;

id Sex; /\* use the re-assgined values as the variable names after transpose \*/

**run**;

**proc** **print** data=want3b;

**run**;

/\* Now add a new obs on top and create a new variable 'Variable': \*/

**data** want3 (keep=Variable Overall Female Male);

length Variable $**20**;

/\* if the first obs, do this, otherwise don't.

- So this block is done only once! \*/

if \_n\_ = **1** then

do; /\* do this for the first obs of this datastep only! \*/

Variable = "Age, year";

output; /\* other three variables will have missing value \*/

end;

/\* after the first obs, will do this: \*/

set want3b;

if strip(V0)="meanSD" then Variable = " Mean(SD)";

else if strip(V0)="medianMM" then Variable = " Median(Min, Max)";

output; /\* need explicit 'output' here because the explicit 'output' before

already overwrote the implicit 'output' \*/

**run**;

**proc** **print** data=want3 noobs style(data)={asis=on};

**run**;