# **/\* The SAS program (HEI-2010 at the population/group level using ASA24-2016 and ASA24-2018 data)**

# **Population Ratio Method.SAS \*/**

/\*This program creates component and total scores of the HEI-2010 for a population or a group. The 12 components include: Total Fruit, Whole Fruit, Total Vegetables, Greens and Beans, Whole Grains, Dairy, Total Protein Foods, Seafood and Plant Proteins, Fatty Acids Refined Grains, Sodium, and Empty Calories. \*/

/\*Please see accompanying readme file. \*/

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**/\*INSTRUCTIONS – complete tasks 1-4 in this section, and run these SAS codes before proceeding to the HEI-2010 scoring program that follows\*/**

/\*1. Create a folder on your computer “home folder”, and save the ASA24 data, and the required HEI-2010 macros in it. Specify the path to the folder. \*/

%let home = C:\Users\Documents\ASA24; /\*In this Example, the “home” folder is in C Drive, within Documents, and is called ASA24\*/

/\*2. Filename here specifies the input file. \*/

filename Totals “&home\Totals\Totals.csv”; /\* In this example, the ASA24-2016 or ASA24-2018 Daily Total Nutrient and Pyramid Equivalents data “Totals”, are saved in a folder called “Totals”, within the “home” folder. The data are in csv format. \*/

/\*3. Create a folder within the “home” folder, where the output file, containing HEI-2010 component and total scores are to be exported. Specify the name of the folder. \*/

filename RES “&home\RES”; /\*In this Example, the folder is called “RES”, within the “home” folder, and the exported results will be a csv file called “result”. \*/

/\*4. Read in required HEI-2010 scoring macros. These macros must be saved within the home folder. \*/

%include “&home\hei2010.beanspeas.allocation.macro.sas';

%include “&home\hei2010.score.macro.sas”;

/\*NOTE: Once you have completed all the steps above, all you need to do is run the SAS program below. Unless you used different names for your folders and datasets, no other action is required from you. \*/

TITLE 'ASA24-2016 and ASA24-2018 HEI-2010 population ratio scores';

/\*Section (I): Calculations at the individual participant level to obtain variables needed to calculate HEI-2010 scores.\*/

/\*Step 1. Locate the required datasets and variables and make necessary edits to the datasets. \*/

\*a. Read in the TOTALS output files from ASA24-2016;

**Proc** **import** datafile=Totals

Out=Totals

Dbms=csv

Replace;

Getnames=yes;

**Run**;

\*Step 2. Create additional required variables;

**DATA** Totals;

SET Totals;

WHOLEFRT=F\_CITMLB+F\_OTHER;

MONOPOLY=MFAT+PFAT;

ALLMEAT=PF\_MPS\_TOTAL+PF\_EGGS+PF\_NUTSDS+PF\_SOY;

SEAPLANT=PF\_SEAFD\_HI+PF\_SEAFD\_LOW+PF\_SOY+PF\_NUTSDS;

/\*\*Calculate intake of Calories from SoFAAS\*\*/

ADDSUGC=**16**\*ADD\_SUGARS; /\*calories from added sugars\*/

SOLFATC=**9**\*SOLID\_FATS; /\*calories from solid fat\*/

maxalcgr=**13**\*(kcal/**1000**); /\*max grams of alcohol based on kcal intake\*/

if ALC <= maxalcgr then EXALCCAL=**0**; /\*consumed less than max\*/

else if ALC > maxalcgr then EXALCCAL=**7**\*(ALC-maxalcgr); /\*get cal from extra alc grams\*/

EMPTYCAL10=ADDSUGC+SOLFATC+EXALCCAL; /\*total empty calories in hei2010 definition\*/

**Run**;

/\*Section (II): Calculation of weighted means and a variance-covariance matrix and generation of a Monte Carlo dataset, enabling standard errors to be calculated. \*/

/\*Step 1. Calculate the weighted means and the variance/covariance matrix for the dietary variables of interest. \*/

**proc** **corr** data=Totals outp=covdata cov nocorr nomiss noprint;

var KCAL V\_TOTAL V\_Drkgr v\_legumes F\_TOTAL WHOLEFRT G\_WHOLE d\_total

allmeat seaplant monopoly sfat SODI G\_REFINED EMPTYCAL10;

**run**;

**proc** **print** data=covdata noobs;

title2 "Cov, Mean, Std, and N from Corr Procedure.";

**run**;

**data** count (keep=\_TYPE\_ sampsize);

set covdata;

if \_TYPE\_='N';

sampsize=KCAL;

\_TYPE\_='COV';

**run**;

**data** mean;

set covdata;

if \_TYPE\_='MEAN';

**run**;

**proc** **sort** data=count;

by \_type\_;

**run**;

**proc** **sort** data=covdata;

by \_type\_;

**run**;

**data** covdata;

merge count covdata;

by \_type\_;

**run**;

**data** covdata;

set covdata;

array diag KCAL V\_TOTAL V\_Drkgr v\_legumes F\_TOTAL WHOLEFRT G\_WHOLE d\_total

allmeat seaplant monopoly sfat SODI G\_REFINED EMPTYCAL10;

do over diag;

diag=diag/sampsize;

end;

if \_TYPE\_ ne 'COV' then delete;

**run**;

**data** covdata;

set covdata mean;

**run**;

**proc** **print**;

title2 'input to simnormal';

**run**;

/\*Step 2. In this step, a Monte Carlo data set with 10,001 rows is generated using the means and variance/covariance matrix from step 1. \*/

**proc** **simnormal** data=covdata(type=cov) numreal=**10000** seed=**51230077** outseed out=sim\_data;

var KCAL V\_TOTAL V\_Drkgr v\_legumes F\_TOTAL WHOLEFRT G\_WHOLE d\_total

allmeat seaplant monopoly sfat SODI G\_REFINED EMPTYCAL10;

**run**;

**proc** **means** data=sim\_data n nmiss min max mean stddev;

var KCAL V\_TOTAL V\_Drkgr v\_legumes F\_TOTAL WHOLEFRT G\_WHOLE d\_total

allmeat seaplant monopoly sfat SODI G\_REFINED EMPTYCAL10;

title2 "Distributions of Simulated Data";

**run**;

**proc** **print** data=sim\_data(obs=**20**);

title2 "Listing of 20 Records from Simulated Data";

**run**;

/\*Section (III): Allocation of beans and peas (legumes) to Total Proteins and Seafood and Plant Proteins

and/or Total Vegetables and Greens and Beans and application of the HEI-2010 scoring algorithm.\*/

/\*Step 1. Allocate beans and peas (legumes) using the Monte Carlo data set from Section II and the beans and peas allocation macro. \*/

%***LEG2010A*** (indat=sim\_data,

kcal=KCAL,

allmeat=allmeat,

seaplant=seaplant,

v\_total=v\_total,

v\_drkgr=v\_drkgr,

legumes=v\_legumes,

outdat=lsim\_data);

/\*Step 2. Apply the HEI-2010 scoring macro.\*/

%***HEI2010*** (indat=lsim\_data,

kcal=KCAL,

lv\_total=legume\_added\_V\_TOTAL,

lbeangrn=legume\_added\_BEANGRN,

f\_total=f\_total,

wholefrt=wholefrt,

g\_whl=G\_WHOLE,

d\_total=d\_total,

lallmeat=legume\_added\_ALLMEAT,

lseaplant=legume\_added\_SEAPLANT,

monopoly=monopoly,

sfat=sfat,

sodi=SODI,

G\_NWHL=G\_REFINED,

EMPTYCAL10=EMPTYCAL10,

outdat=aftermac);

**run**;

**proc** **means** data=aftermac n nmiss min max mean stddev;

var legume\_added\_v\_total legume\_added\_BEANGRN legume\_added\_allmeat legume\_added\_SEAPLANT

vegden grbnden frtden whfrden wgrnden dairyden

meatden seaplden faratio sodden rgden sofa\_perc;

title2 'after legume allocation and hei 2010 scoring macro';

**run**;

/\*Section (IV): Calculation of mean HEI-2010 component and total scores and their confidence intervals.\*/

/\*Step 1. This step uses univariate and means procedures to compute total and component scores and their standard errors. \*/

**proc** **univariate** data=aftermac noprint;

var HEIX1\_TOTALVEG HEIX2\_GREEN\_AND\_BEAN HEIX3\_TOTALFRUIT HEIX4\_WHOLEFRUIT HEIX5\_WHOLEGRAIN HEIX6\_TOTALDAIRY

HEIX7\_TOTPROT HEIX8\_SEAPLANT\_PROT HEIX9\_FATTYACID HEIX10\_SODIUM HEIX11\_REFINEDGRAIN HEIX12\_SOFAAS HEI2010\_TOTAL\_SCORE;

output out=ci pctlpts=**2.5** **97.5** pctlpre=h1\_ h2\_ h3\_ h4\_ h5\_ h6\_ h7\_ h8\_ h9\_ h10\_ h11\_ h12\_ totscore\_;

**run**;

**proc** **means** data=aftermac noprint;

var HEIX1\_TOTALVEG HEIX2\_GREEN\_AND\_BEAN HEIX3\_TOTALFRUIT HEIX4\_WHOLEFRUIT HEIX5\_WHOLEGRAIN HEIX6\_TOTALDAIRY

HEIX7\_TOTPROT HEIX8\_SEAPLANT\_PROT HEIX9\_FATTYACID HEIX10\_SODIUM HEIX11\_REFINEDGRAIN HEIX12\_SOFAAS HEI2010\_TOTAL\_SCORE;

output out=stat min=h1\_min h2\_min h3\_min h4\_min h5\_min h6\_min h7\_min h8\_min h9\_min h10\_min h11\_min h12\_min totscore\_min

max=h1\_max h2\_max h3\_max h4\_max h5\_max h6\_max h7\_max h8\_max h9\_max h10\_max h11\_max h12\_max totscore\_max

mean=h1\_mean h2\_mean h3\_mean h4\_mean h5\_mean h6\_mean h7\_mean h8\_mean h9\_mean h10\_mean h11\_mean h12\_mean totscore\_mean

stddev= h1\_stddev h2\_stddev h3\_stddev h4\_stddev h5\_stddev h6\_stddev h7\_stddev h8\_stddev h9\_stddev h10\_stddev h11\_stddev h12\_stddev totscore\_stddev;

**run**;

/\*Step 2. This step prepares the results for display\*/

**data** all; merge ci stat;

**run**;

**data** result (keep=score slabel min max mean stderr low high);

set all;

score='HEIx1 ';

slabel='HEI-2010 COMPONENT 1 TOTAL VEGETABLES';

min=h1\_min;

max=h1\_max;

mean=h1\_mean;

stderr=h1\_stddev;

low=h1\_2\_5;

high=h1\_97\_5;

output result;

score='HEIx2';

slabel='HEI-2010 COMPONENT 2 GREENS AND BEANS';

min=h2\_min;

max=h2\_max;

mean=h2\_mean;

stderr=h2\_stddev;

low=h2\_2\_5;

high=h2\_97\_5;

output result;

score='HEIx3';

slabel='HEI-2010 COMPONENT 3 TOTAL FRUIT';

min=h3\_min;

max=h3\_max;

mean=h3\_mean;

stderr=h3\_stddev;

low=h3\_2\_5;

high=h3\_97\_5;

output result;

score='HEIx4';

slabel='HEI-2010 COMPONENT 4 WHOLE FRUIT';

min=h4\_min;

max=h4\_max;

mean=h4\_mean;

stderr=h4\_stddev;

low=h4\_2\_5;

high=h4\_97\_5;

output result;

score='HEIx5';

slabel='HEI-2010 COMPONENT 5 WHOLE GRAINS';

min=h5\_min;

max=h5\_max;

mean=h5\_mean;

stderr=h5\_stddev;

low=h5\_2\_5;

high=h5\_97\_5;

output result;

score='HEIx6';

slabel='HEI-2010 COMPONENT 6 DAIRY';

min=h6\_min;

max=h6\_max;

mean=h6\_mean;

stderr=h6\_stddev;

low=h6\_2\_5;

high=h6\_97\_5;

output result;

score='HEIx7';

slabel='HEI-2010 COMPONENT 7 TOTAL PROTEIN FOODS';

min=h7\_min;

max=h7\_max;

mean=h7\_mean;

stderr=h7\_stddev;

low=h7\_2\_5;

high=h7\_97\_5;

output result;

score='HEIx8';

slabel='HEI-2010 COMPONENT 8 SEAFOOD AND PLANT PROTEIN';

min=h8\_min;

max=h8\_max;

mean=h8\_mean;

stderr=h8\_stddev;

low=h8\_2\_5;

high=h8\_97\_5;

output result;

score='HEIx9';

slabel='HEI-2010 COMPONENT 9 FATTY ACID RATIO';

min=h9\_min;

max=h9\_max;

mean=h9\_mean;

stderr=h9\_stddev;

low=h9\_2\_5;

high=h9\_97\_5;

output result;

score='HEIx10';

slabel='HEI-2010 COMPONENT 10 SODIUM';

min=h10\_min;

max=h10\_max;

mean=h10\_mean;

stderr=h10\_stddev;

low=h10\_2\_5;

high=h10\_97\_5;

output result;

score='HEIx11';

slabel='HEI-2010 COMPONENT 11 REFINED GRAINS';

min=h11\_min;

max=h11\_max;

mean=h11\_mean;

stderr=h11\_stddev;

low=h11\_2\_5;

high=h11\_97\_5;

output result;

score='HEIx12';

slabel='HEI-2010 COMPONENT 12 SOFAAS CALORIES';

min=h12\_min;

max=h12\_max;

mean=h12\_mean;

stderr=h12\_stddev;

low=h12\_2\_5;

high=h12\_97\_5;

output result;

score='TOTAL HEI 2010';

slabel='TOTAL HEI-2010 SCORE';

min=totscore\_min;

max=totscore\_max;

mean=totscore\_mean;

stderr=totscore\_stddev;

low=totscore\_2\_5;

high=totscore\_97\_5;

output result;

**run**;

/\*Step 3. This step displays the results\*/

**proc** **print** data=result;

id score;

var slabel min max mean stderr low high;

title2 'simple unweighted survey design population method - mean and confidense interval of HEI 2010 using ASA24-2016 data';

**run**;

/\*Step 4. This step provides an option to export the results into a CSV file that can be opened in Excel. \*/

**proc** **export** data=result

file=res

dbms=csv

replace;

**run**;