# /\*The SAS program (HEI-2010 Individual Scores using multiple recalls from NHANES 2009-2010 data (and FPED))

# ByPerson.SAS\*/

**/\*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 FPED data, NHANES data, Demographic data, and the required HEI-2010 macros in it. Specify the path to the folder. \*/

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

/\*2. Libnames here specify the input files. \*/

libname NH “&home\NH”;

libname FPED “&home\FPED”; /\*In this Example, the FPED data are in a folder called “FPED”, and the NHANES and Demographic data are in a folder called “NH”, all saved within the “home” folder. These are SAS datasets. \*/

/\*3. Create a folder in the "home" folder, where the output file, containing HEI-2010 component and total scores for each respondent 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 “hei2010r”. \*/

/\*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 dataset and folders, no other action is required from you. \*/

title 'HEI-2010 scores for NHANES 2009-2010 day 1, AGE >= 2, RELIABLE DIETS, Include Pregnant and Lactating Women';

/\*Step 1: locate the required datasets and variables \*/

\*part a: get FPED data per day;

**data** FPED1 (keep=SEQN DAYREC F\_CITMLB F\_OTHER PF\_MPS\_TOTAL PF\_EGGS

PF\_NUTSDS PF\_SOY PF\_SEAFD\_HI PF\_SEAFD\_LOW ADD\_SUGARS SOLID\_FATS

V\_TOTAL V\_DRKGR V\_LEGUMES F\_TOTAL G\_WHOLE D\_TOTAL G\_REFINED);

set FPED.fped\_dr1tot\_0910;

DAYREC=1;

F\_CITMLB=DR1T\_F\_CITMLB;

F\_OTHER=DR1T\_F\_OTHER;

PF\_MPS\_TOTAL=DR1T\_PF\_MPS\_TOTAL;

PF\_EGGS=DR1T\_PF\_EGGS;

PF\_NUTSDS=DR1T\_PF\_NUTSDS;

PF\_SOY=DR1T\_PF\_SOY;

PF\_SEAFD\_HI=DR1T\_PF\_SEAFD\_HI;

PF\_SEAFD\_LOW=DR1T\_PF\_SEAFD\_LOW;

ADD\_SUGARS=DR1T\_ADD\_SUGARS;

SOLID\_FATS=DR1T\_SOLID\_FATS;

V\_TOTAL=DR1T\_V\_TOTAL;

V\_DRKGR=DR1T\_V\_DRKGR;

V\_LEGUMES=DR1T\_V\_LEGUMES;

F\_TOTAL=DR1T\_F\_TOTAL;

G\_WHOLE=DR1T\_G\_WHOLE;

D\_TOTAL=DR1T\_D\_TOTAL;

G\_REFINED=DR1T\_G\_REFINED;

run;

**data** FPED2 (keep=SEQN DAYREC F\_CITMLB F\_OTHER PF\_MPS\_TOTAL PF\_EGGS

PF\_NUTSDS PF\_SOY PF\_SEAFD\_HI PF\_SEAFD\_LOW ADD\_SUGARS SOLID\_FATS

V\_TOTAL V\_DRKGR V\_LEGUMES F\_TOTAL G\_WHOLE D\_TOTAL G\_REFINED);

set FPED.fped\_dr2tot\_0910;

DAYREC=2;

F\_CITMLB=DR2T\_F\_CITMLB;

F\_OTHER=DR2T\_F\_OTHER;

PF\_MPS\_TOTAL=DR2T\_PF\_MPS\_TOTAL;

PF\_EGGS=DR2T\_PF\_EGGS;

PF\_NUTSDS=DR2T\_PF\_NUTSDS;

PF\_SOY=DR2T\_PF\_SOY;

PF\_SEAFD\_HI=DR2T\_PF\_SEAFD\_HI;

PF\_SEAFD\_LOW=DR2T\_PF\_SEAFD\_LOW;

ADD\_SUGARS=DR2T\_ADD\_SUGARS;

SOLID\_FATS=DR2T\_SOLID\_FATS;

V\_TOTAL=DR2T\_V\_TOTAL;

V\_DRKGR=DR2T\_V\_DRKGR;

V\_LEGUMES=DR2T\_V\_LEGUMES;

F\_TOTAL=DR2T\_F\_TOTAL;

G\_WHOLE=DR2T\_G\_WHOLE;

D\_TOTAL=DR2T\_D\_TOTAL;

G\_REFINED=DR2T\_G\_REFINED;

run;

data FPED;

set FPED1 FPED2;

run;

\*part b: get individual total food intake if reliable recall status;

**data** NUTRIENT1 (keep= SEQN WTDRD1 KCAL SFAT ALCO SODI DR1DRSTZ MFAT

PFAT DAYREC);

set NH.DR1TOT\_F;

if DR1DRSTZ=**1**; /\*reliable dietary recall status\*/

KCAL=DR1TKCAL;

MFAT=DR1TMFAT;

PFAT=DR1TPFAT;

SFAT=DR1TSFAT;

SODI=DR1TSODI;

ALCO=DR1TALCO;

DAYREC=1;

**run**;

data NUTRIENT2 (keep= SEQN KCAL SFAT ALCO SODI DR2DRSTZ MFAT

PFAT DAYREC);

set NH.DR2TOT\_F;

if DR2DRSTZ=1; /\*reliable dietary recall status \*/

KCAL=DR2TKCAL;

MFAT=DR2TMFAT;

PFAT=DR2TPFAT;

SFAT=DR2TSFAT;

SODI=DR2TSODI;

ALCO=DR2TALCO;

DAYREC=2;

run;

data NUTRIENT;

set NUTRIENT1 NUTRIENT2;

run;

\*part c: get demographic data for persons aged two and older;

**data** DEMO (keep=SEQN RIDAGEYR RIAGENDR SDDSRVYR SDMVPSU SDMVSTRA);

set NH.DEMO\_F;

if RIDAGEYR >= **2**;

**run**;

/\*Step 2: Combine the required datasets\*/

**proc sort data**=FPED;

by SEQN DAYREC;

**run**;

**proc sort data**=NUTRIENT;

by SEQN DAYREC;

**run**;

data NUTFDPYR;

merge FPED (in=F) NUTRIENT (in=N);

by SEQN DAYREC;

if F and N;

run;

**proc sort data**=DEMO;

by SEQN;

**run**;

data COHORT;

merge DEMO (in=C) NUTFDPYR (in=P);

by SEQN;

if C and P;

run;

/\*Step 3: Creates additional required variables: WHOLEFRT, MONOPOLY, ALLMEAT, SEAPLANT, ADDSUGC, SOLFATC, EXALCCAL, and EMPTYCAL10. \*/

**data** COHORT;

set COHORT;

by SEQN;

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\_NUTSDS+PF\_SOY;

/\*\*Calculate intake of Empty Calories\*\*/

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

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

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

if ALCO <= MAXALCGR then EXALCCAL=**0**; /\*consumed less than max\*/

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

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

**run**;

\*get sum per person across all days of variables of interest;

proc means data= COHORT noprint;

by SEQN;

var KCAL ALLMEAT SEAPLANT V\_TOTAL V\_DRKGR V\_LEGUMES F\_TOTAL WHOLEFRT G\_WHOLE D\_TOTAL

MONOPOLY SFAT SODI G\_REFINED EMPTYCAL10;

output out=BYID sum=;

run;

/\*Step 4: Run the macro to properly allocate the intakes of the FPED variable Legumes in cup equivalents (V\_LEGUMES) to either Total Protein Foods and Seafood and Plant Proteins (ALLMEAT and SEAPLANT) or Total Vegetables and Greens and Beans (V\_TOTAL and V\_DRKGR). The four resulting variables from this step, named legume\_added\_V\_TOTAL, legume\_added\_BEANGRN, legume\_added\_ALLMEAT, and legume\_added\_SEAPLANT are all used in the HEI scoring macro in Step 5\*/

%***LEG2010A*** (indat= BYID,

kcal= KCAL,

allmeat=ALLMEAT,

seaplant=SEAPLANT,

v\_total=V\_TOTAL,

v\_drkgr= V\_DRKGR,

legumes= V\_LEGUMES,

outdat=AFTERLEG);

/\*Step 5: Apply the HEI-2010 scoring macro. \*/

%***HEI2010*** (indat= AFTERLEG,

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=HEI2010);

/\*Step 6: Displays and saves the results. \*/

\*part a: this program saves one HEI-2010 score for each individual, based on multiple 24HRs;

**data** HEI2010R (keep=SEQN KCAL 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);

set HEI2010;

**run**;

\*part b: calculates an unweighted mean across all individuals in group;

**proc** **means** n nmiss min max mean data=HEI2010R;

**run**;

\*part c: saves results as CSV file one line per subject;

**proc** **export** data=HEI2010R

file=RES

dbms=csv

replace;

**run**;