

HINTS 4 Cycle 1 History Document

Iteration One (finalized on July 25, 2012)

Data Editing

The following continuous variables were identified to contain invalid or unusual values. Those values were replaced with negative value of -4, "Unreadable or Non-conforming numeric response".

Age: K1 (J1). What is your age?

One respondent had age of 169, which was replaced by -4.

Height_Feet: D13. About how tall are you without shoes? Feet:

One respondent had the feet of 8 and inches of 9 for height. Both feet and inches were replaced by -4.

Height_Inches: D13. About how tall are you without shoes? Inches:

One respondent had the inches of 69 and feet of 5 for height. Both feet and inches were replaced by -4.

Weight: D14. About how much do you weigh, in pounds?

One respondent had weight of 10; one respondent had weight of 43; one respondent had weight of 47; one respondent had weight of 49. All their weights were replaced by -4.

Standard Recode

Standard recode/derived variables are listed below.

AgeGrpA: -->AgeGrpA. (Age Recode -- 4 Levels)

The Age variable was re-coded into 4 categories: 18-34; 35-39; 40-44; 45+. The original negative values were carried over.

AgeGrpB: -->AgeGrpB. (Age Recode -- 5 Levels)

The Age variable was re-coded into 5 categories: 18-34; 35-49; 50-64; 65-74; 75+. The original negative values were carried over.

Educa:-->Educa. What is the highest level of school you completed? (Education Recode -- 4 Levels)

The Education variable was re-coded into 4 categories: Less than High School; High School Graduate; Some College; College Graduate or More. The original negative values were carried over.

EducB:-->EducB. What is the highest level of school you completed? (Education Recode -- 5 Levels)

The Education variable was re-coded into 5 categories: Less than High School; High School Graduate; Some College; Bachelor's Degree; Post-Baccalaureate Degree. The original negative values were carried over.

RaceEthn:-->Race/Ethnicity. (Hispanic, AmerInd, Asian, Black, PacIsl and White Recode -- 7 Levels)

The RaceEthn was created with Hispanic, AmerInd, Asian, Black, PacIsl and White variables. The RaceEthn has 7 categories: Hispanic; Non-Hispanic White; Non-Hispanic Black or African American; Non-Hispanic American Indian or Alaska Native; Non-Hispanic Asian; Non-Hispanic Native Hawaiian or other Pacific Islander; Non-Hispanic Multiple Races Mentioned. If AmerInd, Asian, Black, PacIsl and White all have value of -9, "Missing data (Not Ascertained)", the RaceEthn was assigned with value of -9 if Hispanic=2. The RaceEthn was assigned with value of -9 if Hispanic=-9.

HHInc:-->HHInc. What is your {combined} annual household income? (Incomeranges Recode -- 5 Levels)

The IncomeRanges variable was re-coded into 5 categories: Less than \$20,000; \$20,000 to < \$35,000; \$35,000 to < \$50,000; \$50,000 to < \$75,000; \$75,000 or more. The original negative values were carried over.

Pap3Yr:-->Pap3Yr. When was your last Pap Screen? Guidelines based on recommendations from the US Preventive Services Task Force (WhenPapTest Recode -- 3 Levels)

The WhenPapTest variable was re-coded into 3 categories: Had Pap Screen within last 3 years; Had Pap Screen more than 3 years ago; Never Had Pap Screen. The original negative values were carried over.

BMI:-->BMI. Body Mass Index (Weight*703)/(Height in inches2)**

The BMI variable was created with weight in pounds and height in inches. If feet of height, inches of height, or weight had value of -9, "Missing data (Not Ascertained)", the BMI was assigned to -9. If feet

of height, inches of height, or weight had value of -4, “Unreadable or Non-conforming numeric response”, the BMI was assigned to -4. For BMI variable, the value of -9 was assigned before value of -4.

AgeDX:-->AgeDX. At what age were you diagnosed with cancer? (WhenDiagnosedCancer Recode)

The variable AgeDX is a copy of variable WhenDiagnosedCancer.

TimeSinceDX:-->TimeSinceDX. How long ago were you diagnosed with cancer? (WhenDiagnosedCancer Age Recode)

The variable TimeSinceDX was created with EverHadCancer, WhenDiagnosedCancer and Age variables. The variable TimeSinceDX has 4 categories: Less than 1 Year since DX; 2-5 Years since DX; 6-10 Years since DX; 11+ Years since DX. When the negative values were encountered, the proper negative values were assigned to TimeSinceDX variable.

smokeStat:-->smokeStat. Smoking Status (Smoke100 and SmokeNow Recode)

The variable smokeStat was created with Smoke100 and SmokeNow variables. The variable smokeStat has 3 categories: Current; Former; Never. When the negative values were encountered, the proper negative values were assigned to smokeStat variable.

PHQ4:-->PHQ4. PHQ-4 total score (LittleInterest Hopeless Nervous Worrying Combined)

The variable PHQ4 was created by summing the reversed LittleInterest, Hopeless, Nervous and Worrying variables. When the negative values were encountered, the proper negative values were assigned to PHQ4 variable.

Variable Exclusion

The imputed variables and imputation flag variables created by Westat were excluded from the public data file. They are: WhereSeekHealthInfo_IMP, WhereSeekHealthInfo_IFlag, StrongNeedHealthInfo_IMP, StrongNeedHealthInfo_IFlag, HCCoverage_I, HCCover_IFlag, GenderC_I, Gender_IFlag, EverHadCancer_I, Cancer_IFlag, Age_I, Age_IFlag, MaritalStatus_I, Marital_IFlag, Education_I, Education_IFlag, Hispanic_I, Hisp_IFlag, Race_cat, Race_Cat_I and Race_IFlag.

Format Editing

Formats Added for Existing Variables

The format CENSREG was created and assigned to the variable CENSREG. The format \$CENSIV was created and assigned to the variable CENSIV.

Formats Added for Standard Recode Variables

The formats AgeGrpA, AgeGrpB, EducA, EducB, RaceEthn, HHInc, Pap3Yr, BMI, AgeDX, TimeSinceDX, smokeStat and PHQ4Total were created and assigned to the variables AgeGrpA, AgeGrpB, EducA, EducB, RaceEthn, HHInc, Pap3Yr, BMI, AgeDX, TimeSinceDX, smokeStat and PHQ4, respectively.

Formats Modified for Certain Variables

The formats for certain variables were modified, where the formats of value ranges were removed in order to access the raw values. The modified formats are: MailHHA, \$WhereSG, \$StrongN, \$InternI, \$InternH, \$OthIntF, \$HCCover, HowLonF, HowLong, Average, HeightF, Height_, Weight, DrinksP, SleepWG, SleepWo, SleepWF, SleepWe, \$CaOthF, WhenDia, \$RxDrugH, \$MPRecal, \$FDASitI, Age, \$Occupat, MultiOc, YearCam, TotalHo, HHAdulR, HHAdulI, HHAdulJ, HHAdulK, HHAdulL, Childre, MailSuF and MailSur.

Formats Disassociated for Excluded Variables

The variables excluded from the public data file were also removed from the format statement.

Imputation of Income Variable

The income variable (IncomeRanges) has relatively higher percentage (10.2% for un-weighted percentage or 9.8% for weighted percentage) of missing values. This variable was imputed via PROC HOTDECK in SUDAAN. The imputation class variables are: Education (K5), RaceEthn (standard recode), RentOrOwn (K14), ComfortableEnglish (K8), and BornInUSA (K6). The imputed values were saved in a new variable IncomeRanges_IMP.

SAS Code for Data Editing

```
data HINTS4CYCLE1;
    set HINTS4CYCLE1;

* Re-code the Age of 169 to Unreadable or Non-conforming numeric response
*;
    if Age in (169) then
        Age = -4;

* Re-code the Height_Feet of 8 to Unreadable or Non-conforming numeric
response
*;
    if Height_Feet in (8) then
        do;
            Height_Feet = -4;
            Height_Inches = -4;
        end;

* Re-code the Height_Inches of 69 to Unreadable or Non-conforming numeric
response
*;
    if Height_Inches in (69) then
        do;
            Height_Feet = -4;
            Height_Inches = -4;
        end;

* Re-code the Weight of 10, 43, 47, and 49 to Unreadable or Non-conforming
numeric response
*;
    if Weight in (10, 43, 47, 49) then
        Weight = -4;
run;
```

SAS Code for Standard Recode

```
data HINTS4CYCLE1;
  set HINTS4CYCLE1;

  if 18 <= Age <= 34 then
    AgeGrpA = 1;
  else if 35 <= Age <= 39 then
    AgeGrpA = 2;
  else if 40 <= Age <= 44 then
    AgeGrpA = 3;
  else if 45 <= Age then
    AgeGrpA = 4;
  else if Age in (-4, -9) then
    AgeGrpA = Age;
  label AgeGrpA = '-->AgeGrpA. (Age Recode -- 4 Levels)';

  if 18 <= Age <= 34 then
    AgeGrpB = 1;
  else if 35 <= Age <= 49 then
    AgeGrpB = 2;
  else if 50 <= Age <= 64 then
    AgeGrpB = 3;
  else if 65 <= Age <= 74 then
    AgeGrpB = 4;
  else if 75 <= Age then
    AgeGrpB = 5;
  else if Age in (-4, -9) then
    AgeGrpB = Age;
  label AgeGrpB = '-->AgeGrpB. (Age Recode -- 5 Levels)';

  if Education in (1, 2) then
    EducA = 1;
  else if Education in (3) then
    EducA = 2;
  else if Education in (4, 5) then
    EducA = 3;
  else if Education in (6, 7) then
    EducA = 4;
  else if Education in (-9) then
    EducA = Education;
  label EducA = '-->EducA. What is the highest level of school you
  completed? (Education Recode -- 4 Levels)';

  if Education in (1, 2) then
    EducB = 1;
  else if Education in (3) then
    EducB = 2;
  else if Education in (4, 5) then
    EducB = 3;
  else if Education in (6) then
    EducB = 4;
  else if Education in (7) then
    EducB = 5;
```

```

else if Education in (-9) then
    EducB = Education;
label EducB = '-->EducB. What is the highest level of school you
completed? (Education Recode -- 5 Levels)';

if Hispanic = 1 then
    RaceEthn = 1;
else if Hispanic = 2 then
    do;
        if White = 1 and AmerInd = 2 and Asian = 2 and Black = 2
and PacIsl = 2 then
            RaceEthn = 2;
        else if Black = 1 and AmerInd = 2 and Asian = 2 and PacIsl
= 2 and White = 2 then
            RaceEthn = 3;
        else if AmerInd = 1 and Asian = 2 and Black = 2 and PacIsl
= 2 and White = 2 then
            RaceEthn = 4;
        else if Asian = 1 and AmerInd = 2 and Black = 2 and PacIsl
= 2 and White = 2 then
            RaceEthn = 5;
        else if PacIsl = 1 and AmerInd = 2 and Asian = 2 and Black
= 2 and White = 2 then
            RaceEthn = 6;
        else if AmerInd = -9 and Asian = -9 and Black = -9 and
PacIsl = -9 and White = -9 then
            RaceEthn = -9;
        else
            RaceEthn = 7;
    end;
else if Hispanic = -9 then
    RaceEthn = -9;
label RaceEthn = '-->Race/Ethnicity. (Hispanic, AmerInd, Asian, Black,
PacIsl and White Recode -- 7 Levels)';

if IncomeRanges in (1, 2, 3) then
    HHInc = 1;
else if IncomeRanges in (4) then
    HHInc = 2;
else if IncomeRanges in (5) then
    HHInc = 3;
else if IncomeRanges in (6) then
    HHInc = 4;
else if IncomeRanges in (7, 8, 9) then
    HHInc = 5;
else if IncomeRanges in (-5, -9) then
    HHInc = IncomeRanges;
label HHInc = '-->HHInc. What is your {combined} annual household
income? (IncomeRanges Recode -- 5 Levels)';

if WhenPapTest in (1, 2, 3) then
    Pap3Yr = 1;
else if WhenPapTest in (4, 5) then
    Pap3Yr = 2;
else if WhenPapTest in (6) then
    Pap3Yr = 3;
else if WhenPapTest in (-1, -2, -5, -6, -9) then

```

```

        Pap3Yr = WhenPapTest;
        label Pap3Yr = '-->Pap3Yr.  When was your last Pap Screen?  Guidelines
based on recommendations from the US Preventive Services Task Force
(WhenPapTest Recode -- 3 Levels)';

        if Height_Feet >= 0 and Height_Inches >= 0 and Weight > 0 then
            BMI = (Weight * 703) / ((Height_Feet * 12 + Height_Inches)**2);
        else if Height_Feet = -9 or Height_Inches = -9 or Weight = -9 then
            BMI = -9;
        else if Height_Feet = -4 or Height_Inches = -4 or Weight = -4 then
            BMI = -4;
        label BMI = '-->BMI.  Body Mass Index (Weight*703)/(Height in
inches**2)';
        if BMI not in (-4, -9) then
            BMI = round(BMI, 0.1);

        AgeDX = WhenDiagnosedCancer;
        label AgeDX = '-->AgeDX.  At what age were you diagnosed with cancer?
(WhenDiagnosedCancer Recode)';

        if EverHadCancer in (1) then
            do;
                if Age in (-4, -9) then
                    TimeSinceDX = Age;
                else if WhenDiagnosedCancer in (-1, -2, -6, -9) and Age >=
18 then
                    TimeSinceDX = WhenDiagnosedCancer;
                else if WhenDiagnosedCancer >= 0 and Age >= 18 then
                    do;
                        if 0 <= (Age - WhenDiagnosedCancer) <= 1 then
                            TimeSinceDX = 1;
                        else if 2 <= (Age - WhenDiagnosedCancer) <= 5
then
                            TimeSinceDX = 2;
                        else if 6 <= (Age - WhenDiagnosedCancer) <= 10
then
                            TimeSinceDX = 3;
                        else if 11 <= (Age - WhenDiagnosedCancer) then
                            TimeSinceDX = 4;
                    end;
                end;
            else if EverHadCancer in (-9) then
                TimeSinceDX = EverHadCancer;
            else if EverHadCancer in (2) then
                TimeSinceDX = WhenDiagnosedCancer;
            label TimeSinceDX = '-->TimeSinceDX.  How long ago were you diagnosed
with cancer?  (WhenDiagnosedCancer Age Recode)';

        if Smoke100 in (1) then
            do;
                if SmokeNow in (1, 2) then
                    smokeStat = 1;
                else if SmokeNow in (3) then
                    smokeStat = 2;
                else if SmokeNow in (-9) then
                    smokeStat = SmokeNow;
            end;

```

```

else if Smoke100 in (2) then
    do;
        smokeStat = 3;
    end;
else if Smoke100 in (-9) then
    smokeStat = Smoke100;
label smokeStat = '-->smokeStat. Smoking Status (Smoke100 and SmokeNow
Recode)';

array a(*) LittleInterest Hopeless Nervous Worrying;

PHQ4 = 0;
do I = 1 to dim(a);
    if PHQ4 not in (-5, -9) and a(I) in (1, 2, 3, 4) then
        PHQ4 = PHQ4 + (4 - a(I));
    else
        do;
            if PHQ4 not in (-5, -9) and a(I) in (-5, -9) then
                PHQ4 = a(I);
            else
                PHQ4 = PHQ4;
        end;
    end;
label PHQ4 = '-->PHQ4. PHQ-4 total score (LittleInterest Hopeless
Nervous Worrying Combined)';

drop I;

run;

```

SAS Code for Variable Exclusion

```
data HINTS4CYCLE1;
  set HINTS4CYCLE1;

  drop Age_I Age_IFlag
        EverHadCancer_I Cancer_IFlag
        Education_I Education_IFlag
        GenderC_I Gender_IFlag
        HCCoverage_I HCCover_IFlag
        Hispanic_I Hisp_IFlag
        MaritalStatus_I Marital_IFlag
        Race_Cat_I Race_IFlag Race_cat
        StrongNeedHealthInfo_IMP StrongNeedHealthInfo_IFlag
        WhereSeekHealthInfo_IMP WhereSeekHealthInfo_IFlag
  ;

run;
```

SAS Code for Format Editing

SAS Code for Formats Added for Existing Variables

```
value CENSREG 1 = 'NORTHEAST CENSUS REGION'
              2 = 'MIDWEST CENSUS REGION'
              3 = 'SOUTH CENSUS REGION'
              4 = 'WEST CENSUS REGION'
              ;

value $CENSNDIV "00" = 'Pacific Census Division'
              "01" = 'New England Census Division'
              "02" = 'Middle Atlantic Census Division'
              "03" = 'East North Central Census Division'
              "04" = 'West North Central Census Division'
              "05" = 'South Atlantic Census Division'
              "06" = 'East South Central Census Division'
              "07" = 'West South Central Census Division'
              "08" = 'Mountain Census Division'
              "09" = 'Pacific Census Division'
              ;
```

SAS Code for Formats Added for Standard Recode Variables

```
value AgeGrpA 1 = '18-34'
              2 = '35-39'
              3 = '40-44'
              4 = '45+'
              -4 = 'Unreadable or Nonconforming Numeric Response'
              -9 = 'Missing Data (Not Ascertained)'
              ;

value AgeGrpB 1 = '18-34'
              2 = '35-49'
              3 = '50-64'
              4 = '65-74'
              5 = '75+'
              -4 = 'Unreadable or Nonconforming Numeric Response'
              -9 = 'Missing Data (Not Ascertained)'
              ;

value EducA 1 = 'Less than High School'
            2 = 'High School Graduate'
            3 = 'Some College'
            4 = 'College Graduate or More'
            -9 = 'Missing Data (Not Ascertained)'
            ;

value EducB 1 = 'Less than High School'
            2 = 'High School Graduate'
            3 = 'Some College'
```

```

    4 = "Bachelor's Degree"
    5 = 'Post-Baccalaureate Degree'
    -9 = 'Missing Data (Not Ascertained)'
;

value RaceEthn 1 = 'Hispanic'
    2 = 'Non-Hispanic White'
    3 = 'Non-Hispanic Black or African American'
    4 = 'Non-Hispanic American Indian or Alaska
Native'
    5 = 'Non-Hispanic Asian'
    6 = 'Non-Hispanic Native Hawaiian or other
Pacific Islander'
    7 = 'Non-Hispanic Multiple Races Mentioned'
    -9 = 'Missing Data (Not Ascertained)'
;

value HHInc 1 = 'Less than $20,000'
    2 = '$20,000 to < $35,000'
    3 = '$35,000 to < $50,000'
    4 = '$50,000 to < $75,000'
    5 = '$75,000 or More'
    -5 = 'Multiple Responses Selected in Error'
    -9 = 'Missing Data (Not Ascertained)'
;

value Pap3Yr 1 = 'Had Pap Screen within last 3 yrs'
    2 = 'Had Pap Screen more than 3 yrs ago'
    3 = 'Never Had Pap Screen'
    -1 = 'Inapplicable, coded 1 in GenderC'
    -2 = 'Question Answered in Error (Commission Error)'
    -5 = 'Multiple Responses Selected in Error'
    -6 = 'Missing Data (Filter Missing)'
    -9 = 'Missing Data (Not Ascertained)'
;

value BMI -4 = 'Unreadable or Nonconforming Numeric Response'
    -9 = 'Missing Data (Not Ascertained)'
;

value AgeDX -1 = 'Inapplicable, coded 0 in EverHadCancer'
    -2 = 'Question Answered in Error (Commission Error)'
    -6 = 'Missing Data (Filter Missing)'
    -9 = 'Missing Data (Not Ascertained)'
;

value TimeSinceDX 1 = 'Less than 1 Yr Since DX'
    2 = '2-5 Yrs Since DX'
    3 = '6-10 Yrs Since DX'
    4 = '11+ Yrs Since DX'
    -1 = 'Inapplicable, coded 0 in EverHadCancer'
    -2 = 'Question Answered in Error (Commission
Error)'
    -4 = 'Unreadable or Nonconforming Numeric
Response'
    -9 = 'Missing Data (Not Ascertained)'
;

```

```

value smokeStat 1 = 'Current'
                2 = 'Former'
                3 = 'Never'
                -9 = 'Missing Data (Not Ascertained)'
;

value PHQ4Total -5 = 'Multiple Responses Selected in Error'
               -9 = 'Missing Data (Not Ascertained)'
;

```

SAS Code for Formats Modified for Certain Variables

The deleted formats for value regions were strikethrough and highlighted with gray color.

```

Value MailHHA
-1 = "Inapplicable, coded 0 in AdultsInHH"
-2 = "Question answered in error (Commission Error)"
-6 = "Missing data (Filter Missing)"
-9 = "Missing data (Not Ascertained)"
01-05 = "Number of adults living in the household"
;

Value $WhereSG
"-1" = "Inapplicable, coded 0 in SeekHealthInfo, or coded 1 - 12
in WhereSeekHealthInfo"
"-2" = "Question answered in error (Commission Error)"
"-6" = "Missing data (Filter Missing)"
"-9" = "Missing data (Not Ascertained)"
"A"-"Z" = "Verbatim response"
;

Value $StrongN
"-1" = "Inapplicable, coded 1 - 12 in StrongNeedHealthInfo"
"-2" = "Question answered in error (Commission Error)"
"-6" = "Missing data (Filter Missing)"
"-9" = "Missing data (Not Ascertained)"
"A"-"Z" = "Verbatim response"
;

Value $InternI
"-1" = "Inapplicable, coded 0 in UseInternet or coded 0 in
Internet_Other"
"-2" = "Question answered in error (Commission Error)"
"-6" = "Missing data (Filter Missing)"
"-9" = "Missing data (Not Ascertained)"
"A"-"Z" = "Verbatim response"
;

```

```

Value $InternH
  "-1" = "Inapplicable, coded 0 in UseInternet or coded 0 in
InternetHealthInfoSite"
  "-2" = "Question answered in error (Commission Error)"
  "-6" = "Missing data (Filter Missing)"
  "-9" = "Missing data (Not Ascertained)"
"A" - "Z" = "Verbatim response"
;

Value $OthIntF
  "-1" = "Inapplicable, coded 0 in UseInternet or this is a Short
Form household or coded 0 in"
  "-2" = "Question answered in error (Commission Error)"
  "-6" = "Missing data (Filter Missing)"
  "-9" = "Missing data (Not Ascertained)"
"A" - "Z" = "Verbatim response"
;

Value $HCCover
  "-1" = "Inapplicable, coded 0 in HCCoverage_Other"
  "-6" = "Missing data (Filter Missing)"
  "-9" = "Missing data (Not Ascertained)"
"A" - "Z" = "Verbatim response"
;

Value HowLonF
  -1 = "Inapplicable, coded 0 in TimesModerateExercise"
  -2 = "Question answered in error (Commission Error)"
  -4 = "Unreadable or Non-conforming numeric response"
  -6 = "Missing data (Filter Missing)"
  -9 = "Missing data (Not Ascertained)"
00-90 = "Length of exercise time, in minutes"
;

Value HowLong
  -1 = "Inapplicable, coded 0 in TimesModerateExercise"
  -2 = "Question answered in error (Commission Error)"
  -4 = "Unreadable or Non-conforming numeric response"
  -6 = "Missing data (Filter Missing)"
  -9 = "Missing data (Not Ascertained)"
00-06 = "Length of exercise time, in hours"
;

Value Average

```

```
-4 = "Unreadable or Non-conforming numeric response"  
-9 = "Missing data (Not Ascertained)"  
00-24 = "Length of TV/Game time, in hours"  
;
```

```
Value HeightF  
-4 = "Unreadable or Non-conforming numeric response"  
-9 = "Missing data (Not Ascertained)"  
4-7 = "Height, in feet"  
;
```

```
Value Height_  
-4 = "Unreadable or Non-conforming numeric response"  
-9 = "Missing data (Not Ascertained)"  
00-12 = "Height, in inches"  
;
```

```
Value Weight  
-4 = "Unreadable or Non-conforming numeric response"  
-9 = "Missing data (Not Ascertained)"  
75-600 = "Weight, in pounds"  
;
```

```
Value DrinksP  
-1 = "Inapplicable, coded 0 in DrinksPerWeek"  
-2 = "Question answered in error (Commission Error)"  
-4 = "Unreadable or Non-conforming numeric response"  
-6 = "Missing data (Filter Missing)"  
-9 = "Missing data (Not Ascertained)"  
01-20 = "Average number of drinks per day"  
;
```

```
Value SleepWG  
-4 = "Unreadable or Non-conforming numeric response"  
-9 = "Missing data (Not Ascertained)"  
01-16 = "Hours of sleep, workday"  
;
```

```
Value SleepWo  
-4 = "Unreadable or Non-conforming numeric response"  
-9 = "Missing data (Not Ascertained)"  
00-60 = "Minutes of sleep, workday"  
;
```

```

Value SleepWF
  -4 = "Unreadable or Non-conforming numeric response"
  -9 = "Missing data (Not Ascertained)"
  01-16 = "Hours of sleep, weekend"
;

Value SleepWe
  -4 = "Unreadable or Non-conforming numeric response"
  -9 = "Missing data (Not Ascertained)"
  00-60 = "Minutes of sleep, weekend"
;

Value $CaOtherF
CaOther"
  "-1" = "Inapplicable, coded 0 in EverHadCancer or coded 0 in
  "-2" = "Question answered in error (Commission Error)"
  "-6" = "Missing data (Filter Missing)"
  "-9" = "Missing data (Not Ascertained)"
  "A"-"Z" = "Verbatim response"
;

Value WhenDia
  -1 = "Inapplicable, coded 0 in EverHadCancer"
  -2 = "Question answered in error (Commission Error)"
  -6 = "Missing data (Filter Missing)"
  -9 = "Missing data (Not Ascertained)"
  00-97 = "Age at diagnosis"
;

Value $RxDrugH
in BuyRxDrugs or coded 0"
  "-1" = "Inapplicable, this is a Short Form household or coded 0
  "-2" = "Question answered in error (Commission Error)"
  "-6" = "Missing data (Filter Missing)"
  "-9" = "Missing data (Not Ascertained)"
  "A"-"Z" = "Verbatim response"
;

Value $MPRecal
in MPRecall_Other"
  "-1" = "Inapplicable, this is a Short Form household or coded 0
  "-2" = "Question answered in error (Commission Error)"
  "-6" = "Missing data (Filter Missing)"

```

```

"-9" = "Missing data (Not Ascertained)"
"A" - "Z" = "Verbatim response"
;

Value $FDASitI
"-1" = "Inapplicable, this is a Short Form household or coded 1
in VisitFDAWebsite or coded 0"
"-2" = "Question answered in error (Commission Error)"
"-6" = "Missing data (Filter Missing)"
"-9" = "Missing data (Not Ascertained)"
"A" - "Z" = "Verbatim response"
;

Value Age
-4 = "Unreadable or Non-conforming numeric response"
-9 = "Missing data (Not Ascertained)"
18-100 = "Current age, in years"
;

Value $Occupat
"-1" = "Inapplicable, coded 1 - 6 in OccupationStatus"
"-6" = "Missing data (Filter Missing)"
"-9" = "Missing data (Not Ascertained)"
"A" - "Z" = "Verbatim response"
;

Value MultiOc
0-7 = "total responses checked in question K2 (J2)"
;

Value YearCam
-1 = "Inapplicable, coded 1 in BornInUSA"
-2 = "Question answered in error (Commission Error)"
-6 = "Missing data (Filter Missing)"
-9 = "Missing data (Not Ascertained)"
1911-2011 = "Year"
;

Value TotalHo
-9 = "Missing data (Not Ascertained)"
01-10 = "Number of people"
;

```

```
Value HHAdulR
  -4 = "Unreadable or Non-conforming numeric response"
  -9 = "Missing Data"
  18-99 = "Age, in years"
;
```

```
Value HHAdulI
  -1 = "Innapplicable, Fewer than 2 adults in HH"
  -4 = "Unreadable or Non-conforming numeric response"
  -6 = "Missing data (Filter Missing)"
  -9 = "Missing Data"
  18-99 = "Age, in years"
;
```

```
Value HHAdulJ
  -1 = "Innapplicable, Fewer than 3 adults in HH"
  -4 = "Unreadable or Non-conforming numeric response"
  -6 = "Missing data (Filter Missing)"
  -9 = "Missing Data"
  18-99 = "Age, in years"
;
```

```
Value HHAdulK
  -1 = "Innapplicable, Fewer than 4 adults in HH"
  -4 = "Unreadable or Non-conforming numeric response"
  -6 = "Missing data (Filter Missing)"
  -9 = "Missing Data"
  18-99 = "Age, in years"
;
```

```
Value HHAdulL
  -1 = "Innapplicable, Fewer than 5 adults in HH"
  -4 = "Unreadable or Non-conforming numeric response"
  -6 = "Missing data (Filter Missing)"
  -9 = "Missing Data"
  18-99 = "Age, in years"
;
```

```
Value Childre
  -4 = "Unreadable or Non-conforming numeric response"
  -9 = "Missing Data"
  00-05 = "Number of children under 18"
;
```

```
Value MailSuF
  -1 = "Innapplicable, time recorded in "Hours" only"
  -4 = "Unreadable or Non-conforming numeric response"
  -9 = "Missing Data"
  00-60 = "Minutes"
;
```

```
Value MailSur
  -1 = "Innapplicable, time recorded in "Minutes" only"
  -4 = "Unreadable or Non-conforming numeric response"
  -9 = "Missing Data"
  00-04 = "Hours"
;
```

SAS Code for Imputation of Income Variable

```
data HINTS4CYCLE1;
  set HINTS4CYCLE1;

  COPY_Education = Education;
  if COPY_Education in (-9) then
    COPY_Education = .;

  COPY_RaceEthn = RaceEthn;
  if COPY_RaceEthn in (-9) then
    COPY_RaceEthn = .;

  COPY_RentOrOwn = RentOrOwn;
  if COPY_RentOrOwn in (-5, -9) then
    COPY_RentOrOwn = .;

  COPY_ComfortableEnglish = ComfortableEnglish;
  if COPY_ComfortableEnglish in (-5, -9) then
    COPY_ComfortableEnglish = .;

  COPY_BornInUSA = BornInUSA;
  if COPY_BornInUSA in (-9) then
    COPY_BornInUSA = .;

  COPY_IncomeRanges = IncomeRanges;
  if COPY_IncomeRanges in (-5, -9) then
    COPY_IncomeRanges = .;

  format COPY_Education Educati. COPY_RaceEthn RaceEthn. COPY_RentOrOwn
RentOrO.
          COPY_ComfortableEnglish Comfort. COPY_BornInUSA BornInU.;
run;

proc freq data=HINTS4CYCLE1;
  tables COPY_Education*Education / missing;
  tables COPY_RaceEthn*RaceEthn / missing;
  tables COPY_RentOrOwn*RentOrOwn / missing;
  tables COPY_ComfortableEnglish*ComfortableEnglish / missing;
  tables COPY_BornInUSA*BornInUSA / missing;
  tables COPY_IncomeRanges*IncomeRanges / missing;
run;

data HINTS4CYCLE1;
  set HINTS4CYCLE1;

  ID = _N_;
run;

proc sort data=HINTS4CYCLE1;
  by COPY_Education COPY_RaceEthn COPY_RentOrOwn COPY_ComfortableEnglish
COPY_BornInUSA;
run;

proc hotdeck data=HINTS4CYCLE1;
```

```

weight person_finwt0;
impvar COPY_IncomeRanges;
impby COPY_Education COPY_RaceEthn COPY_RentOrOwn
COPY_ComfortableEnglish COPY_BornInUSA;
impname COPY_IncomeRanges="IncomeRanges_IMP";
impid ID;
output IMPID IMPBY IMPUTEVAL / filename=imputel replace;
run;

proc freq data=imputel;
tables IncomeRanges_IMP / missing;
run;

proc contents data=imputel;
run;

proc sort data=HINTS4CYCLE1;
by ID;
run;

proc sort data=imputel (keep=ID IncomeRanges_IMP);
by ID;
run;

data HINTS4CYCLE1;
merge HINTS4CYCLE1 (in=A) imputel (in=B);
by ID;

if A = 1 and B = 1;
run;

data _null_;
set HINTS4CYCLE1;

if IncomeRanges not in (-5, -9) and COPY_IncomeRanges ^=
IncomeRanges_IMP then
put ID IncomeRanges COPY_IncomeRanges IncomeRanges_IMP;
run;

data HINTS4CYCLE1;
set HINTS4CYCLE1;

if missing(IncomeRanges_IMP) = 1 then
IncomeRanges_IMP = IncomeRanges;
label IncomeRanges_IMP = '-->IncomeRanges_IMP. Imputed IncomeRanges
variable via PROC HOTDECK in SUDAAN';

drop COPY_Education COPY_RaceEthn COPY_RentOrOwn
COPY_ComfortableEnglish COPY_BornInUSA
ID FIPS FIPST
COPY_IncomeRanges;

format _all_;
run;

```

