

HINTS FDA, Cycle 2 History Document

November 2017

Data Editing

The following variables were identified to contain invalid or unusual values. Those values were replaced with negative value of -4, “Unreadable or Non-conforming numeric response”, negative value of -9, “Missing data (Not Ascertained)” or reasonable regular values.

MailHHAdults: 2. Including yourself, how many people age 18 or older live in this household?
Adultsinhh: 1. Is there more than one person age 18 or older living in this household?

Eighty-seven respondents had value 1 for both MailHHAdults and Adultsinhh, which was invalid. Adultsinhh value for these 87 records was replaced by 2 and MailHHAdults value was set to -2.

R_HHAdults: Reconciled number of adults in household

Three respondents had missing values (.), which were replaced by -9.

Totalhousehold: G12. Including yourself, how many people live in your household?

Twenty-nine respondents had Totalhousehold of 0, which were replaced by 1.

Standard Recode

Standard recode/derived variables are listed below.

SEC_RUCA_2010_DESCRIPT: USDA 2010 Secondary Rural-Urban Commuting Area Codes
Description (variable found in Stata dataset, only)

The SEC_RUCA_2010 variable was recoded into the string variable SEC_RUCA_2010_DESCRIPT containing the secondary urban-rural commuting area code description. This recoded variable is present on the Stata dataset only, as Stata does not allow format/labels to be applied to the non-integer values that are present in SEC_RUCA_2010.

AgeGrpB: AgeGrpB. 5 Level Age Categories Version B (Derived from Age)

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.

AgeGrpC: AgeGrpC. 5 Level Age Categories Version C (Derived from Age)

The Age variable was re-coded into 5 categories: 18-29; 30-49; 50-64; 65-74; 75+. The original negative values were carried over.

EducA: EducA. What is the highest level of school you completed? 5 Levels (Derived from Education)

The Education variable was re-coded into 5 categories: Less than High School; High School Graduate; Post High School Training other than College (vocational or technical); Some College; College Graduate or More. The original negative values were carried over.

RaceEthn: Race/Ethnicity. 7 Levels (Derived from Hisp_Cat and Race_Cat2)

The RaceEthn was created with Hisp_Cat and Race_Cat2 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 Hisp_Cat had a value of 10, "Not Hispanic", and Race_Cat2 had value of -9, "Missing data (Not Ascertained)", the RaceEthn was assigned with value of -9. The RaceEthn was assigned with value of -9 if Hispanic=-9.

RaceEthn5: Race/Ethnicity. 5 Levels (Derived from Hisp_Cat and Race_Cat)

The RaceEthn5 was created with Hisp_Cat and Race_Cat2 variables. The RaceEthn5 has 5 categories: Non-Hispanic White; Non-Hispanic Black or African American; Hispanic; Non-Hispanic Asian; Non-Hispanic Other. If Hisp_Cat had value of 10, "Not Hispanic", and Race_Cat2 had value of -9, "Missing data (Not Ascertained)", the RaceEthn5 was assigned with value of -9. The RaceEthn5 was assigned with value of -9 if Hispanic=-9.

HHInc: HHInc. What is your {combined} annual household income? 5 Levels (Derived from IncomeRanges Recode)

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.

SmokeStat: SmokeStat. Smoking Status (Derived from Smoke100 and SmokeNow)

The variable smokeStat was created with Smoke100 and SmokeNow variables. The variable smokeStat has 3 categories: Current; Former; Never. If Smoke100 had value of 1 and SmokeNow had value of -5, "Multiple responses selected in error", the smokeStat was assigned to -4. If Smoke100 had value of 1 and SmokeNow had value of -9, "Missing data (Not Ascertained)", the smokeStat was assigned to -9. If Smoke100 had value of -9, "Missing data (Not Ascertained)", the smokeStat was assigned to -6.

SmokelessStat: SmokelessStat. Smokeless tobacco use status (Derived from UsedTobacco20Times and UseTobaccoNow)

The variable SmokelessStat was created with UsedTobacco20Times and UseTobaccoNow variables. The variable SmokelessStat has 3 categories: Current; Former; Never. If UsedTobacco20Times had value of 1 and UseTobaccoNow had value of -5, "Multiple responses selected in error", the SmokelessStat was assigned to -4. If UsedTobacco20Times had value of 1 and UseTobaccoNow had value of -9, "Missing data (Not Ascertained)", the SmokelessStat was assigned to -9. If UsedTobacco20Times had value of -9, "Missing data (Not Ascertained)", the SmokelessStat was assigned to -6.

CigarStat: CigarStat. Cigar Smoking status (Derived from NumberCigarsSmoked and SmokeNowCigars)

The variable CigarStat was created with NumberCigarsSmoked and SmokeNowCigars variables. The variable CigarStat has 3 categories: Current; Former; Never. If NumberCigarsSmoked had value of 4 or 5 and SmokeNowCigars had value of -5, "Multiple responses selected in error", the CigarStat was assigned to -4. If NumberCigarsSmoked had value of 4 or 5 and SmokeNowCigars had value of -9, "Missing data (Not Ascertained)", the CigarStat was assigned to -9. If NumberCigarsSmoked had value of -9, "Missing data (Not Ascertained)", the CigarStat was assigned to -6.

eCigUse: eCigUse. Electronic Cigarette Use (Derived from UsedECigEver and UseECigNow)

The variable eCigUse was created with UsedECigEver and UseECigNow variables. The variable eCigUse has 3 categories: Current; Former; Never. If UsedECigEver had value of 1 and UseECigNow had value of -5, "Multiple responses selected in error", the eCigUse was assigned to -4. If UsedECigEver had value of 1 and UseECigNow had value of -9, "Missing data (Not Ascertained)", the eCigUse was assigned to -9. If UsedECigEver had value of -9, "Missing data (Not Ascertained)", the eCigUse was assigned to -6.

PolyUserA: PolyuserA. PolyTobacco Status - Indicator of current polytobacco use (Derived from SmokeStat [cigarettes], SmokelessStat [smokeless tobacco], and CigarStat [cigars], only)

The variable PolyUserA was created with smokeStat, SmokelessStat, and CigarStat. If more than one of these variables were equal to 1 (Current users) then PolyUserA was set to 1 (current polytobacco user). If two or more of these variables were equal to 2 (Former user) or 3 (Never user) then PolyUserA was set to 0 (not current polytobacco user). Else enough data was not available to make a polytobacco user determination and PolyUserA was set to -9 "Missing data (Not Ascertained)".

PolyUserB: PolyuserB. PolyTobacco Status - Indicator of current polytobacco use (Derived from SmokeStat [cigarettes], SmokelessStat [smokeless tobacco], CigarStat [cigars], and eCigUse [e-cigarettes])

The variable PolyUserB was created with smokeStat, SmokelessStat, CigarStat, and eCigUse. If more than one of these variables were equal to 1 (Current users) then PolyUserB was set to 1 (current polytobacco user). If three or more of these variables were equal to 2 (Former user) or 3 (Never user) then PolyUserB was set to 0 (not current polytobacco user). Else enough data was not available to make a polytobacco user determination and PolyUserB was set to -9 "Missing data (Not Ascertained)".

Label Editing

Labels Added for Standard Recode Variables

Labels were created for the following recoded variables: SEC_RUCA_2010_DESCRIPT, AgeGrpB, AgeGrpC, EducA, RaceEthn, RaceEthn5, HHInc, SmokeStat, SmokelessStat, CigarStat, eCigUse, PolyUserA, and PolyUserB.

Labels Modified for Certain Variables

Labels were modified for the following variables: NCHSURCODE2013, PR_RUCA_2010, SEC_RUCA_2010, and NicotineWantSmoke.

Format Editing

Formats Added for Standard Recode Variables

The formats AgeGrpB, AgeGrpC, EducA, RaceEthn, RaceEthnN, HHInc, smokeStat, ecigStat, and polyuser were created and assigned to the variables AgeGrpB, AgeGrpC, EducA, RaceEthn, RaceEthn5, HHInc, SmokeStat, SmokelessStat (smokeStat format), CigarStat (smokeStat format), eCigUse, PolyUserA, and PolyUserB respectively.

Formats Modified for Certain Variables

All skip patterns in formats were modified (i.e. all instructions to skip questions were deleted). The modified formats are: ADULTSI, ANYONER, AWAREQU, BORNINU, CONSID, ECIG_OT, ECIGINF, LOOKEDE, NOTICEH, OCCUPAF, RECENTF, SEEKHEA, SEXUALF, SMOKENF, SMOKENO, TOBACCL, TOBACCN, TOBACCP, TOBACCO, USEDECI, and USEINTE.

The format SEC_RUC was modified to contain the full USDA 2010 Secondary Rural-Urban Commuting Area descriptions to avoid truncation.

Imputation of Income Variable

The income variable (IncomeRanges) has relatively higher percentage (11% for un-weighted percentage or 9% for weighted percentage) of missing values. This variable was imputed via PROC IMPUTE in SUDAAN. The imputation class variables are: Education (G6), RaceEthn (standard recode), RentOrOwn (G15), BornInUSA (G7) and SpeakEnglish (G9). Since the variable SpeakEnglish was asked for people who were born outside USA (BornInUSA=2), the variable SpeakEnglish was declared after BornInUSA in imputation class statement. The copy variables of the imputation class variables and income variable were created, where the missing values were appropriately coded. The copy variables are used for the imputation. The imputed values were saved in a new variable IncomeRanges_IMP.

SAS Code for Data Editing

```
* Recode R_HHAdults of missing (.) to -9
*;
    if missing(R_HHAdults) = 1 then
        R_HHAdults = -9;

* Recode Totalhousehold of 0 to 1
*;
    if Totalhousehold=0 then Totalhousehold=1;

* Recode MailHHAdults of 0 to 1
*;
    if MailHHAdults=0 then
        MailHHAdults = 1;

* Recode adultsinhh and MailhhAdults;
    if adultsinhh=1 and MailHHAdults=1 Then do;
```

```
adulttsinhh=2;  
MailHHadults=-2;
```

```
end;
```

SAS Code for Standard Recode

```
if sec_ruca_2010 = 1 then sec_ruca_2010_descript = 'Metropolitan area  
core: primary flow within an urbanized area (UA), No additional code';  
else if sec_ruca_2010 = 1.1 then sec_ruca_2010_descript = 'Metropolitan  
area core: primary flow within an urbanized area (UA), Secondary flow 30% to  
50% to a larger UA';  
else if sec_ruca_2010 = 2 then sec_ruca_2010_descript = 'Metropolitan  
area high commuting: primary flow 30% or more to a UA, No additional code';  
else if sec_ruca_2010 = 2.1 then sec_ruca_2010_descript = 'Metropolitan  
area high commuting: primary flow 30% or more to a UA, Secondary flow 30% to  
50% to a larger UA';  
else if sec_ruca_2010 = 3 then sec_ruca_2010_descript = 'Metropolitan  
area low commuting: primary flow 10% to 30% to a UA, No additional code';  
else if sec_ruca_2010 = 4 then sec_ruca_2010_descript = 'Micropolitan  
area core: primary flow within an Urban Cluster of 10,000 to 49,999 (large  
UC), No additional code';  
else if sec_ruca_2010 = 4.1 then sec_ruca_2010_descript = 'Micropolitan  
area core: primary flow within an Urban Cluster of 10,000 to 49,999 (large  
UC), Secondary flow 30% to 50% to a UA';  
else if sec_ruca_2010 = 5 then sec_ruca_2010_descript = 'Micropolitan  
high commuting: primary flow 30% or more to a large UC, No additional code';  
else if sec_ruca_2010 = 5.1 then sec_ruca_2010_descript = 'Micropolitan  
high commuting: primary flow 30% or more to a large UC, Secondary flow 30% to  
50% to a UA';  
else if sec_ruca_2010 = 6 then sec_ruca_2010_descript = 'Micropolitan  
low commuting: primary flow 10% to 30% to a large UC, No additional code';  
else if sec_ruca_2010 = 7 then sec_ruca_2010_descript = 'Small town  
core: primary flow within an Urban Cluster of 2,500 to 9,999 (small UC), No  
additional code';  
else if sec_ruca_2010 = 7.1 then sec_ruca_2010_descript = 'Small town  
core: primary flow within an Urban Cluster of 2,500 to 9,999 (small UC),  
Secondary flow 30% to 50% to a UA';  
else if sec_ruca_2010 = 7.2 then sec_ruca_2010_descript = 'Small town  
core: primary flow within an Urban Cluster of 2,500 to 9,999 (small UC),  
Secondary flow 30% to 50% to a large UC';  
else if sec_ruca_2010 = 8 then sec_ruca_2010_descript = 'Small town  
high commuting: primary flow 30% or more to a small UC, No additional code';  
else if sec_ruca_2010 = 8.1 then sec_ruca_2010_descript = 'Small town  
high commuting: primary flow 30% or more to a small UC, Secondary flow 30% to  
50% to a UA';  
else if sec_ruca_2010 = 8.2 then sec_ruca_2010_descript = 'Small town  
high commuting: primary flow 30% or more to a small UC, Secondary flow 30% to  
50% to a large UC';  
else if sec_ruca_2010 = 9 then sec_ruca_2010_descript = 'Small town low  
commuting: primary flow 10% to 30% to a small UC, No additional code';  
else if sec_ruca_2010 = 10 then sec_ruca_2010_descript = 'Rural areas:  
primary flow to a tract outside a UA or UC, No additional code';
```

```

    else if sec_ruca_2010 = 10.1 then sec_ruca_2010_descript = 'Rural
areas: primary flow to a tract outside a UA or UC, Secondary flow 30% to 50%
to a UA';
    else if sec_ruca_2010 = 10.2 then sec_ruca_2010_descript = 'Rural
areas: primary flow to a tract outside a UA or UC, Secondary flow 30% to 50%
to a large UC';
    else if sec_ruca_2010 = 10.3 then sec_ruca_2010_descript = 'Rural
areas: primary flow to a tract outside a UA or UC, Secondary flow 30% to 50%
to a small UC';
    label sec_ruca_2010_descript = "USDA 2010 Secondary Rural-Urban
Commuting Area Codes Description (variable found in Stata dataset, only)";

```

```

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<0 then
    AgeGrpB = Age;
label AgeGrpB = 'AgeGrpB. 5 Level Age Categories Version B (Derived
from Age; see History Document for more information)';

```

```

if 18 <= Age <= 29 then
    AgeGrpC = 1;
else if 30 <= Age <= 49 then
    AgeGrpC = 2;
else if 50 <= Age <= 64 then
    AgeGrpC = 3;
else if 65 <= Age <= 74 then
    AgeGrpC = 4;
else if 75 <= Age then
    AgeGrpC = 5;
else if Age<0 then
    AgeGrpC = Age;
label AgeGrpC = 'AgeGrpC. 5 Level Age Categories Version C (Derived
from Age; see History Document for more information)';

```

```

if Education in (1, 2) then
    EducA = 1;
else if Education in (3) then
    EducA = 2;
else if Education in (4) then
    EducA = 3;
else if Education in (5) then
    EducA = 4;
else if Education in (6, 7) then
    EducA = 5;
else if Education<0 then
    EducA = Education;

```

```

label EducA = 'EducA. What is the highest level of school you
completed? 5 Levels (Derived from Education; see History Document for more
information)';

```

```

if Hisp_Cat in (21, 22, 23, 24, 25) then
    RaceEthn = 1;
else if Hisp_Cat in (10) then
    do;
        if Race_Cat2 in (11) then
            RaceEthn = 2;
        else if Race_Cat2 in (12) then
            RaceEthn = 3;
        else if Race_Cat2 in (14) then
            RaceEthn = 4;
        else if Race_Cat2 in (31, 32, 33, 34, 35, 36, 37) then
            RaceEthn = 5;
        else if Race_Cat2 in (51, 52, 53, 54) then
            RaceEthn = 6;
        else if Race_Cat2 in (16) then
            RaceEthn = 7;
        else if Race_Cat2 in (-1,-4,-9) then
            RaceEthn = -9;
    end;
else if Hisp_Cat in (-1,-4,-9) then
    do;
        RaceEthn = -9;
    end;

```

```

label RaceEthn = 'Race/Ethnicity. 7 Levels (Derived from Hisp_Cat and
Race_Cat2; see History Document for more information)';

```

```

if Hisp_Cat in (21, 22, 23, 24, 25) then
    RaceEthn5 = 3;
else if Hisp_Cat in (10) then do;
    if Race_Cat2 in (11) then
        RaceEthn5 = 1;
    else if Race_Cat2 in (12) then
        RaceEthn5 = 2;
    else if Race_Cat2 in (31, 32, 33, 34, 35, 36, 37) then
        RaceEthn5 = 4;
    else if Race_Cat2 in (51, 52, 53, 54,14,16) then
        RaceEthn5 = 5;
    else if Race_Cat2 in (-1,-4,-9) then
        RaceEthn5 = -9;
end;
else if Hisp_Cat in (-1,-4,-9) then do;
    RaceEthn5 = -9;
end;

```

```

label RaceEthn5 = 'Race/Ethnicity. 5 Levels (Derived from Hisp_Cat and
Race_Cat2; see History Document for more information)';

```

```

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 (-1,-4,-9) then
        HHInc = IncomeRanges;
    label HHInc = 'HHInc. What is your {combined} annual household income?
5 Levels (Derived from IncomeRanges Recode; see History Document for more
information)';

    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 (-5) then
                smokeStat = -4;
            else if SmokeNow in (-9) then
                smokeStat = -9;
        end;
    else if Smoke100 in (2) then
        do;
            smokeStat = 3;
        end;
    else if Smoke100 in (-9) then
        smokeStat = -6;
    label smokeStat = 'SmokeStat. Smoking Status (Derived from Smoke100
and SmokeNow; see History Document for more information)';

    if UsedTobacco20Times in (1) then do;
        if UseTobaccoNow in (1, 2) then
            smokelessstat = 1;
        else if UseTobaccoNow in (3) then
            smokelessstat = 2;
        else if UseTobaccoNow in (-5) then
            smokelessstat = -4;
        else if UseTobaccoNow in (-9) then
            smokelessstat = -9;
    end;
    else if UsedTobacco20Times in (2) then do;
        smokelessstat = 3;
    end;
    else if UsedTobacco20Times in (-9) then
        smokelessstat = -6;
    label smokelessstat = 'SmokelessStat. Smokeless tobacco use status
(Derived from UsedTobacco20Times and UseTobaccoNow; see History Document for
more information)';

    if NumberCigarsSmoked in (4,5) then do;
        if SmokeNowCigars in (1, 2) then
            CigarStat = 1;
        else if SmokeNowCigars in (3) then
            CigarStat = 2;
        else if SmokeNowCigars in (-5) then
            CigarStat = -4;
        else if SmokeNowCigars in (-9) then
            CigarStat = -9;
    end;

```



```

end;
else if NumberCigarsSmoked in (0,1,2,3) then do;
    CigarStat = 3;
end;
else if NumberCigarsSmoked in (-9) then
    CigarStat = -6;
else if NumberCigarsSmoked in (-5) then
    CigarStat=-4;
label CigarStat = 'CigarStat. Cigar Smoking status (Derived from
NumberCigarsSmoked and SmokeNowCigars; see History Document for more
information)';

if UsedECigEver in (1) then
do;
    if UseECigNow in (1, 2) then
        eCigUse = 1;
    else if UseECigNow in (3) then
        eCigUse = 2;
    else if UseECigNow in (-5) then
        eCigUse = -4;
    else if UseECigNow in (-9) then
        eCigUse = -9;
    end;
else if UsedECigEver in (2) then
do;
    eCigUse = 3;
end;
else if UsedECigEver in (-9) then
    eCigUse = -6;
label eCigUse = 'eCigUse. Electronic Cigarette Use (Derived from
UsedECigEver and UseECigNow; see History Document for more information)';

array inv (4) smokestat smokelessstat cigarstat eciguse;
array cr (4) curr1 curr2 curr3 curr4; /*flags current user*/
array ncr (4) ncurr1 ncurr2 ncurr3 ncurr4; /*flags non-current user*/
do i=1 to 4;
    if inv(i) = 1 then do;
        cr(i)=1;
        ncr(i)=0;
    end;
    else if inv(i)>1 then do;
        cr(i)=0;
        ncr(i)=1;
    end;
end;
if sum(curr1,curr2,curr3)>1 then PolyUserA=1;
else if sum(ncurr1,ncurr2,ncurr3)=>2 then PolyUserA=0;
else PolyUserA=-9;
label polyuserA='PolyuserA. PolyTobacco Status - Indicator of current
polytobacco use (Derived from SmokeStat [cigarettes], SmokelessStat
[smokeless tobacco], and CigarStat [cigars], only; see History Document for
more information)';

if sum(curr1,curr2,curr3,curr4)>1 then PolyUserB=1;
else if sum(ncurr1,ncurr2,ncurr3,ncurr4)=>3 then PolyUserB=0;
else PolyUserB=-9;

```

```

label polyuserB='PolyuserB. PolyTobacco Status - Indicator of current
polytobacco use (Derived from SmokeStat [cigarettes], SmokelessStat
[smokeless tobacco], CigarStat [cigars], and eCigUse [e-cigarettes]; see
History Document for more information)';
drop curr1 curr2 curr3 curr4 ncurr1 ncurr2 ncurr3 ncurr4 i;

```

SAS Code for Format Editing

SAS Code for Formats Added for Standard Recode Variables

```

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 AgeGrpC 1 = '18-29'
2 = '30-49'
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 = 'Vocational or Technical'
4 = 'Some College'
5 = 'College Graduate or More'
-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'
-4 = 'Unreadable or Nonconforming Numeric Response'
-9 = 'Missing Data (Not Ascertained)'

```

```

;

value RaceEthnN
1 = "Non-Hispanic White"
2= "Non-Hispanic Black or African American"
3="Hispanic"
4="Non-Hispanic Asian"
5="Non-Hispanic Other"
-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 smokeStat 1 = 'Current'
2 = 'Former'
3 = 'Never'
-4 = 'Unreadable or Nonconforming Numeric Response'
-6 = 'Missing Data (Filter Missing), coded -9 in Smoke100'
-9 = 'Missing Data (Not Ascertained)'
;

value polyuser 1="Current polytobacco User"
0="Not Current polytobacco User"
-9 = "Missing Data (Not Ascertained)";

value ecigStat
1 = 'Current'
2 = 'Former'
3 = 'Never'
-4 = 'Unreadable or Nonconforming Numeric Response'
-6 = 'Missing Data (Filter Missing), coded -9 in UsedECigEver'
-9 = 'Missing Data (Not Ascertained)';

format AgeGrpC AgeGrpC.
AgeGrpB AgeGrpB.
EducA EducA.
RaceEthn RaceEthn.
RaceEthn5 RaceEthnN.
HHInc HHInc.
SmokeStat smokeStat.
Smokelessstat smokeStat.
CigarStat smokeStat.
ECigUse ECigStat.
polyuserA polyuser.
polyuserB polyuser.
IncomeRanges_imp IncomeR.

```

SAS Code for Imputation of Income Variable

```
* Impute IncomeRanges via PROC HOTDECK
*;
data HINTSFDA2;
    set HINTSFDA2;

    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_SpeakEnglish = SpeakEnglish;
    if COPY_SpeakEnglish in (-1, -2, -5, -6, -9) then
        COPY_SpeakEnglish = .;

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

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

    ID = _N_;

    format COPY_Education Educati. COPY_RaceEthn RaceEthn.
    COPY_RentOrOwn RentOrO.
           COPY_SpeakEnglish SpeakEn. COPY_BornInUSA BornInU.;
run;

proc freq data=HINTSFDA2;
    tables COPY_Education*Education / list missing;
    tables COPY_RaceEthn*RaceEthn / list missing;
    tables COPY_RentOrOwn*RentOrOwn / list missing;
    tables COPY_SpeakEnglish*SpeakEnglish / list missing;
    tables COPY_BornInUSA*BornInUSA / list missing;
    tables COPY_IncomeRanges*IncomeRanges / list missing;
run;

proc impute data= HINTSFDA2 method=wshd notsorted;
    weight person_finwt0;
    impvar COPY_IncomeRanges;
```

```

        imby COPY_Education COPY_RaceEthn COPY_RentOrOwn COPY_BornInUSA
COPY_SpeakEnglish;
        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= HINTSFDA2;
    by ID;
run;

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

data HINTSFDA2;
    merge HINTSFDA2 (in=A) imputel (in=B);
    by ID;

    if A = 1 and B = 1;
run;

data _null_;
    set HINTSFDA2;

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

data HINTSFDA2;
    set HINTSFDA2;

    if missing(IncomeRanges_IMP) = 1 then
        IncomeRanges_IMP = IncomeRanges;
    label IncomeRanges_IMP = ' -->IncomeRanges_IMP. Imputed IncomeRanges
variable via PROC HOTDECK in SUDAAN (Derived variable; see History Document
for more information)';
    format IncomeRanges_IMP IncomeR.;

    drop COPY_Education COPY_RaceEthn COPY_RentOrOwn
COPY_SpeakEnglish COPY_BornInUSA
    ID
    COPY_IncomeRanges;

```

```
format MultiOcc;  
run;
```