National Cancer Institute

TOBACCO USE SUPPLEMENT
CURRENT POPULATION SURVEY

Tobacco Use Supplement to The Current Population Survey: Tips and Tricks of Handling the TUS DATA

James “Todd” Gibson
Information Management Services, Inc.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

Tips and Tricks of Handling the TUS DATA

- Basic information on Getting Started with the Data
- Merging Replicate Weights
- Working with Multiple Years
- Merging Overlap
- Linking to Other CPS Files/Other Supplement to the CPS

Tips and Tricks of Handling the TUS DATA
Basic Information on Getting Started with the Data

- Public Use Files are ASCII Text Files
- Technical Documentation Included With The Data
  - Overview of the Current Populations Survey (CPS)
  - Overview of the Tobacco Use Supplement (TUS)
  - Record Layout of the File
  - TUS Questionnaire
  - Source and Accuracy Statement
  - User Notes for Updates to Files
- Core and Supplement Variables
  - Core: State and Other Geographic Information, Family Income, Race, Origin, Gender, Age, Education, Marital Status, Labor Force Information, Occupation
  - Supplement: Language of Interview, Interview Method (Telephone/In Person), Relationship of Proxy, Cigarette Smoking Prevalence, Smoking History, Menthol Use, Cost of Cigarettes, Use of Other Forms of Tobacco, Smoking Policy at Home, Workplace Smoking Policy, Attitudes Towards Smoking, Medical and Dental Advice to Quit

Tips and Tricks of Handling the TUS DATA
Basic Information on Getting Started with the Data

- Better Estimates Using All 3 Months of Data Collection
- Supplement Weights
  - Non-Response Weights: Analyses that Included Both Self and Proxy Respondents
  - Self-Response Weights: Analyses of Self Respondents Only
- Example Using May 2010, August 2010 and January 2011 Data
  - Variables: Age, Gender, Race/Ethnicity, Education, Interview, PRS64 (Who is Responding to the Supplement), Self Response Weights
  - GESTFIPS: State (FIPS)
  - Selections: PRPERTYP=2, PEAGE>17, INTRVIEW=1, PRS64=1
  - Since Each Survey is Weighted to the Population, Divide Weights by 3
  - Table of Current Smoking Prevalence Rates by Gender and by Age

```
/*****************************************************************/
/* Program: webinar.example.sas                                  */
/*    Date: September 2013                                       */
/*****************************************************************/
Filename cpsmay10 "may10pub.dat" lrecl=1636;
Filename cpsaug10 "aug10pub.dat" lrecl=1636;
Filename cpsjan11 "jan11pub.dat" lrecl=1636;

Proc Format;
  Value PESexF
    1 = "Male"
    2 = "Female"
    ;
  Value AgeGrpF
    1 = "18-24"
    2 = "25-44"
    3 = "45-64"
    4 = "65+"
    ;
  Value SmokStF
    1,4 = "Never/Former"
    2,3 = "Current"
    ;
```

```
%Macro ReadData(survey);
  Data CPS&survey;
    Infile CPS&survey;
    Input @0016 HRMonth  2.
          @0018 HRYear4  4.
          @0093 GESTFIPS 2.     /* State (FIPS) Variable */
          @0122 PEAge    2.
          @0129 PESex    2.
          @0137 PEEduca  2.
          @0139 PRDTRace 2.
          @0141 PRDTHSP  2.
          @0161 PRPERTYP 2.
          @0955 PEA1     2.
          @0967 PEA3     2.
          @1575 Intrview 2.
          @1579 SmokStat 2.
          @1585 PRS64    2.
          @1627 PWSRWgt  10.;
    If PRPerTyp=2;     /* Adult Civilian Household Member */
    If PEAge>17;       /* Ages 18 and over                */
    If Intrview=1;     /* Supplement Interview            */
    If PRS64=1;        /* Self Respondents                */
    If PEA1 In (1,2) & PEA3 Not In (-9,-3,-2); /* Exclude DK,Ref,NR */
    PWSRWgt=PWSRWgt/10000;
%Mend;

%ReadData(may10); Run;
%ReadData(aug10); Run;
%ReadData(jan11); Run;
```

```
Data CPS1011;
  Set CPSmay10 CPSaug10 CPSjan11;
  PWSRWgt=PWSRWgt/3;
  If (18<=PEAge<=24) Then AgeGrp=1;
  Else If (25<=PEAge<=44) Then AgeGrp=2;
  Else If (45<=PEAge<=64) Then AgeGrp=3;
  Else If (PEAge>64) Then AgeGrp=4;
  Format AgeGrp AgeGrpF. PESex PESexF. SmokStat SmokStF.;
  Label HRMonth  = "Month of Interview"
        HRYear4  = "Year of Interview"
        GESTFIPS = "Federal Information Processing Standards State Code"
        PEAge    = "Persons Age"
        PESex    = "Gender"
        PEEduca  = "Highest Level of School Completed"
        PRDTRace = "Race"
        PRDTHSP  = "Detailed Hispanic Origin Group"
        PRPERTYP = "Type of Person Record Recode"
        PEA1     = "Smoked at Least 100 Cigarette in Entire Life"
        PEA3 ="Now Smokes Cigarettes Every Day, Some Days or Not at All"
        Intrview = "Interview Status Recode"
        SmokStat = "Smoker Recode"
        PRS64    = "Who is Responding for the Supplement"
        PWSRWgt  = "Self Response Weight"
       AgeGrp   = "Age Group";
Run;
```

```
Proc Tabulate Data=CPS1011;
  Title1 "Tobacco Use Supplement to The Current Population Survey";
  Title2 "May 2010, August 2010 and January 2011 Combined";
  Title3 "Interviewed, Adults, Ages 18 and Over";
  Title4 "Self Respondents Only";
  Title6 "Current Cigarette Smoking Status By Gender and Age Group";
  Class PESex AgeGrp SmokStat;
  Var PWSRWgt;
  Table All="Total" PESex AgeGrp,
    (SmokStat All="Total")*
    (N="Sample"*F=Comma7.
     PWSRWgt=""*(Sum="Population"*F=Comma12.
     PctSum<SmokStat All>="Percent"*F=7.2))/RTS=11;
Run;
```

Tobacco Use Supplement to The Current Population Survey
May 2010, August 2010 and January 2011 Combined
Interviewed, Adults, Ages 18 and Over
Self Respondents Only

Current Cigarette Smoking Status By Gender and Age Group

Self Respondents Only	Smoker Recode Never/Former Sample	Smoker Recode Never/Former Population	Smoker Recode Never/Former Percent	Smoker Recode Current Sample	Smoker Recode Current Population	Smoker Recode Current Percent	Total Sample	Total Population	Total Percent
Total	142,858	191,855,795	83.92	27,611	36,750,328	16.08	170,469	228,606,123	100.00
Gender Male	61,700	90,423,112	81.95	13,643	19,912,250	18.05	75,343	110,335,362	100.00
Gender Female	81,158	101,432,683	85.76	13,968	16,838,078	14.24	95,126	118,270,761	100.00
Age Group 18-24	11,205	24,568,709	82.91	2,445	5,065,624	17.09	13,650	29,634,333	100.00
Age Group 25-44	47,525	66,086,762	82.06	10,844	14,449,364	17.94	58,369	80,536,127	100.00
Age Group 45-64	52,511	65,621,718	82.19	11,561	14,217,318	17.81	64,072	79,839,036	100.00
Age Group 65+	31,617	35,578,605	92.18	2,761	3,018,023	7.82	34,378	38,596,627	100.00

Tips and Tricks of Handling the TUS DATA
Merging Replicate Weights

- Two Replicate Weight Files For Each Survey in 1992-2011Weights
  - Non-Response Weights
  - Self-Response Weights
- 48 Replicate Weights in 1992-1993 Files
- 80 Replicate Weights in 1995-2003 Files
- 160 Replicate Weights in 2006-2011 Files
- Unique Identifiers
  - 1992-1993: Household ID (H_ID) and Persons' Line Number (A_LINENO)
  - 1995-1996, 1998-1999: Household ID (HRHHID), Serial Suffix (HRSERSUF), Persons' Line Number (PULINENO)
  - 2001-02, 2003, 2006-07, 2010-11: Unique Household Identifier (QSTNUMBER), Unique Person Identifier (OCCURNUM)

Tips and Tricks of Handling the TUS DATA Merging Replicate Weights

- Example Merging Replicate Weights and Calculating Current Smoking Prevalence Using SAS and SUDAAN Using the May 2010, August 2010 and January 2011 Data.
  - Read in Main Survey Data
  - Read In Replicate Weight Data. Multiple Lines Per Record.
  - Sort and Merge Each Survey By QSTNUM and OCCURNUM
  - Divide weights by 3
  - Using Replicate Weights in SUDAAN to Calculate Current Smoking Prevalence, Standard Errors and 95% Confidence Intervals
    - Design=BRR (Balance Repeated Replication)
    - ADJFAY=4
  - Table of Current Smoking Prevalence Rates by Gender

```
/*****************************************************************/
/* Program: webinar.example.reps.sas                             */
/*    Date: September 2013                                       */
/*****************************************************************/
Filename cpsmay10 "may10pub.dat" lrecl=1636;
Filename cpsaug10 "aug10pub.dat" lrecl=1636;
Filename cpsjan11 "jan11pub.dat" lrecl=1636;
Filename repmay10 "may10srrep.dat";
Filename repaug10 "aug10srrep.dat";
Filename repjan11 "jan11srrep.dat";

Proc Format;
  Value PESexF
    1 = "Male"
    2 = "Female"
    ;
  Value AgeGrpF
    1 = "18-24"
    2 = "25-44"
    3 = "45-64"
    4 = "65+"
    ;
  Value CurrentF
    1 = "Yes"
    2 = "No"
    ;
```

```
%Macro ReadData(survey);
  Data CPS&survey;
    Infile CPS&survey;
    Input @0016 HRMonth  2.
          @0018 HRYear4  4.
          @0093 GESTFIPS 2.     /* State (FIPS) Variable */
          @0122 PEAge    2.
          @0129 PESex    2.
          @0161 PRPERTYP 2.
          @0815 QstNum   5.
          @0820 OccurNum 2.
          @0955 PEA1     2.
          @0967 PEA3     2.
          @1575 Intrview 2.
          @1579 SmokStat 2.
          @1585 PRS64    2.
          @1627 PWSRWgt  10.;
    If PRPerTyp=2;     /* Adult Civilian Household Member */
    If PEAge>17;       /* Ages 18 and over                */
    If Intrview=1;     /* Supplement Interview            */
    If PRS64=1;        /* Self Respondents                */

  Data Rep&Survey;
    Infile Rep&Survey Missover;
    Input @001 QstNum   5.
          @006 OccurNum 2.
          @045 SmplWgt  11.4
          @056 (RepWt001-RepWt004) (11.4)
         / @001 (RepWt005-RepWt011) (11.4)
         / @001 (RepWt012-RepWt018) (11.4)
         / @001 (RepWt019-RepWt025) (11.4)
         / @001 (RepWt026-RepWt032) (11.4)
         / @001 (RepWt033-RepWt039) (11.4)
         / @001 (RepWt040-RepWt046) (11.4)
```

```
          / @001 (RepWt047-RepWt053) (11.4)
          / @001 (RepWt054-RepWt060) (11.4)
          / @001 (RepWt061-RepWt067) (11.4)
          / @001 (RepWt068-RepWt074) (11.4)
          / @001 (RepWt075-RepWt081) (11.4)
          / @001 (RepWt082-RepWt088) (11.4)
          / @001 (RepWt089-RepWt095) (11.4)
          / @001 (RepWt096-RepWt102) (11.4)
          / @001 (RepWt103-RepWt109) (11.4)
          / @001 (RepWt110-RepWt116) (11.4)
          / @001 (RepWt117-RepWt123) (11.4)
          / @001 (RepWt124-RepWt130) (11.4)
          / @001 (RepWt131-RepWt137) (11.4)
          / @001 (RepWt138-RepWt144) (11.4)
          / @001 (RepWt145-RepWt151) (11.4)
          / @001 (RepWt152-RepWt158) (11.4)
          / @001 (RepWt159-RepWt160) (11.4);
    If (SmplWgt=0) & (RepWt001 In (.,0)) Then Delete;

  Proc Sort Data=CPS&Survey;
    By QstNum OccurNum;

  Proc Sort Data=Rep&Survey;
    By QstNum OccurNum;

  Data CPS&Survey;
    Merge CPS&Survey(In=In1)
          Rep&Survey;
    By QstNum OccurNum;
    If In1;
%Mend;
```

```
%ReadData(may10); Run;
%ReadData(aug10); Run;
%ReadData(jan11); Run;

Data CPS1011;
  Array Wgts(161) SmplWgt RepWt001-RepWt160;
  Set CPSmay10 CPSaug10 CPSjan11;
  PWSRWgt=PWSRWgt/3;
  Do I = 1 to 161;
    Wgts(I)=Wgts(I)/3;
    End;
  If (18<=PEAge<=24) Then AgeGrp=1;
  Else If (25<=PEAge<=44) Then AgeGrp=2;
  Else If (45<=PEAge<=64) Then AgeGrp=3;
  Else If (PEAge>64) Then AgeGrp=4;
  If SmokStat In (1,4) Then Current=2;
  Else If SmokStat In (2,3) Then Current=1;
  Format AgeGrp AgeGrpF. PESex PESexF. Current CurrentF.;
  Label PESex    = "Gender"
        SmokStat = "Smoker Recode"
        AgeGrp   = "Age Group"
        Current  = "Current Cigarette Smoker"
        QstNum   = "Unique Household Identifier"
        OccurNum = "Unique Person Identifier";
Run;
```

```
Proc Crosstab Data=CPS1011 Design=BRR;      /* SUDAAN CrossTab Procedure */
  Setenv DecWidth=4;
  Weight SmplWgt;
  RepWgt RepWt001-RepWt160/ADJFay=4;
  Class Current PESex AgeGrp/Nofreqs;
  Tables PESex*Current AgeGrp*Current;
  RTitle "Tobacco Use Supplement to The Current Population Survey";
  RTitle "May 2010, August 2016 and January 2011";
  RTitle "Interviewed, Adults, Ages 18 and Over";
  RTitle "Self Respondents Only";
  RTitle "Current Cigarette Smoking Status By Gender and Age Group";
  Print NSum = "Sample Size"
        WSum = "Population Size"
        RowPer = "Percent"
        SERow = "Standard Error"
        LowRow = "Lower 95% CI"
        UpRow = "Upper 95% CI"
        /NSumFmt=F8.0 WSumFmt=F10.0;Run;
```

Variance Estimation Method: BRR
Tobacco Use Supplement to The Current Population Survey
May 2010, August 2010 and January 2011
Interviewed, Adults, Ages 18 and Over
Self Respondents Only
Current Cigarette Smoking Status By Gender and Age Group
by: Gender, Current Cigarette Smoker.

Gender	Current Cigarette Smoker Total	Current Cigarette Smoker Yes	Current Cigarette Smoker No
Total Sample Size	170,469	27,611	142,858
Total Population Size	228,606,123	36,750,328	191,855,795
Total Percent	100.0000	16.0758	83.9242
Total Standard Error	0.0000	0.1347	0.1347
Total Lower 95% CI		15.8117	83.6565
Total Upper 95% CI		16.3435	84.1883
Male Sample Size	75,343	13,643	61,700
Male Population Size	110,335,362	19,912,250	90,423,112
Male Percent	100.0000	18.0470	81.9530
Male Standard Error	0.0000	0.1873	0.1873
Male Lower 95% CI		17.6801	81.5801
Male Upper 95% CI		18.4199	82.3199
Female Sample Size	95,126	13,968	81,158
Female Population Size	118,270,761	16,838078	101,432,683
Female Percent	100.0000	14.2369	85.7631
Female Standard Error	0.0000	0.1484	0.1484
Female Lower 95% CI		13.9469	85.4681
Female Upper 95% CI		14.5319	86.0531

Tips and Tricks of Handling the TUS DATA Working With Multiple Years of Data

- Example Merging Replicate Weights and Calculating Current Smoking Prevalence Using SAS and SUDAAN Using Data From 2 Survey Time Periods (2003 & 2006-2007)
  - Read In and Merge Main Survey and Replicate Weights
  - To Combined 2003 and 2006-07 Need to Construct a New Set of 240 Replicate Weights
    - 2003: 80 Replicate Weights
    - 2006-2007: 160 Replicate Weights
    - ADJFAY=16
  - Table of Current Smoking Prevalence Rates by Gender

```
/*****************************************************************/
/* Program: webinar.example.reps.multiple.years.sas              */
/*    Date: September 2013                                       */
/*****************************************************************/
Filename cpsfeb03 "cpsfeb03.dat" lrecl=1384;
Filename cpsjun03 "cpsjun03.dat" lrecl=1384;
Filename cpsnov03 "cpsnov03.dat" lrecl=1384;
Filename cpsmay06 "cpsmay06.dat" lrecl=1313;
Filename cpsaug06 "cpsaug06.dat" lrecl=1313;
Filename cpsjan07 "cpsjan07.dat" lrecl=1313;
Filename repfeb03 "feb03nrrep.fixed_15yr.dat";
Filename repjun03 "jun03nrrep.fixed_15yr.dat";
Filename repnov03 "nov03nrrep.fixed_15yr.dat";
Filename repmay06 "may06nrrep.new.dat";
Filename repaug06 "aug06nrrep.new.dat";
Filename repjan07 "jan07nrrep.new.dat";

Proc Format;
  Value PESexF
    1 = "Male"
    2 = "Female"
    ;
  Value AgeGrpF
    1 = "18-24"
    2 = "25-44"
    3 = "45-64"
    4 = "65+"
    ;
  Value CurrentF
    1 = "Yes"
    2 = "No"
    ;
```

```
%Macro Read2003(survey);
  Data CPS&survey;
    Infile CPS&survey;
    Input @0016 HRMonth  2.
          @0018 HRYear4  4.
          @0093 GESTFIPS 2.
          @0122 PEAge    2.
          @0129 PESex    2.
          @0161 PRPERTYP 2.
          @0815 QstNum   5.
          @0820 OccurNum 2.
          @0879 PEA1     2.
          @0883 PEA3     2.
          @1347 Intrview 2.
          @1351 SmokStat 2.;
    If PRPerTyp=2;     /* Adult Civilian Household Member */
    If PEAge>17;       /* Ages 18 and over                */
    If Intrview=1;     /* Supplement Interview            */
    If PEA1 In (1,2) & PEA3 Not In (-9,-3,-2); /* Excl Don't Know, Ref, No Resp */

  Data Rep&Survey;
    Infile Rep&Survey Missover;
    Input @001 QstNum   5.
          @007 OccurNum 2.
          @045 SmplWgt  11.4
          @056 (RepWt001-RepWt004) (11.4)
        / @001 (RepWt005-RepWt011) (11.4)
        / @001 (RepWt012-RepWt018) (11.4)
        / @001 (RepWt019-RepWt025) (11.4)
        / @001 (RepWt026-RepWt032) (11.4)
        / @001 (RepWt033-RepWt039) (11.4)
        / @001 (RepWt040-RepWt046) (11.4)
        / @001 (RepWt047-RepWt053) (11.4)
        / @001 (RepWt054-RepWt060) (11.4)
        / @001 (RepWt061-RepWt067) (11.4)
```

```
          / @001 (RepWt068-RepWt074) (11.4)
          / @001 (RepWt075-RepWt080) (11.4);
      If SmplWgt=0 & (RepWt001=0 | RepWt001=.) Then Delete;

    Proc Sort Data=CPS&Survey;
      By QstNum OccurNum;

    Proc Sort Data=Rep&Survey;
      By QstNum OccurNum;

    Data CPS&Survey;
      Merge CPS&Survey(In=In1)
            Rep&Survey;
      By QstNum OccurNum;
      If In1;
%Mend;
```

```
%Macro Read0607(survey);
  Data CPS&survey;
    Infile CPS&survey;
    Input @0016 HRMonth  2.
          @0018 HRYear4  4.
          @0093 GESTFIPS 2.     /* State (FIPS) Variable */
          @0122 PEAge    2.
          @0129 PESex    2.
          @0161 PRPERTYP 2.
          @0815 QstNum   5.
          @0820 OccurNum 2.
          @0955 PEA1     2.
          @0959 PEA3     2.
          @1278 Intrview 2.
          @1282 SmokStat 2.
          @1288 PRS64    2.;
    If PRPerTyp=2;     /* Adult Civilian Household Member */
    If PEAge>17;       /* Ages 18 and over                */
    If Intrview=1;     /* Supplement Interview            */
    If PEA1 In (1,2) & PEA3 Not In (-9,-3,-2); /* Exc Don't Know, Ref, No Resp */

 Data Rep&Survey;
    Infile Rep&Survey Missover;
    Input @001 QstNum   5.
          @007 OccurNum 2.
          @045 SmplWgt  11.4
          @056 (RepWt001-RepWt004) (11.4)
         / @001 (RepWt005-RepWt011) (11.4)
         / @001 (RepWt012-RepWt018) (11.4)
         / @001 (RepWt019-RepWt025) (11.4)
         / @001 (RepWt026-RepWt032) (11.4)
         / @001 (RepWt033-RepWt039) (11.4)
         / @001 (RepWt040-RepWt046) (11.4)
         / @001 (RepWt047-RepWt053) (11.4)
         / @001 (RepWt054-RepWt060) (11.4)
```

```
          / @001 (RepWt061-RepWt067) (11.4)
          / @001 (RepWt068-RepWt074) (11.4)
          / @001 (RepWt075-RepWt081) (11.4)
          / @001 (RepWt082-RepWt088) (11.4)
          / @001 (RepWt089-RepWt095) (11.4)
          / @001 (RepWt096-RepWt102) (11.4)
          / @001 (RepWt103-RepWt109) (11.4)
          / @001 (RepWt110-RepWt116) (11.4)
          / @001 (RepWt117-RepWt123) (11.4)
          / @001 (RepWt124-RepWt130) (11.4)
          / @001 (RepWt131-RepWt137) (11.4)
          / @001 (RepWt138-RepWt144) (11.4)
          / @001 (RepWt145-RepWt151) (11.4)
          / @001 (RepWt152-RepWt158) (11.4)
          / @001 (RepWt159-RepWt160) (11.4);
    If SmplWgt=0 & (RepWt001=0 | RepWt001=.) Then Delete;

  Proc Sort Data=CPS&Survey;
    By QstNum OccurNum;

  Proc Sort Data=Rep&Survey;
    By QstNum OccurNum;

  Data CPS&Survey;
    Merge CPS&Survey(In=In1)
          Rep&Survey;
    By QstNum OccurNum;
    If In1;
%Mend;
```

```
%Read2003(feb03); Run;
%Read2003(jun03); Run;
%Read2003(nov03); Run;
%Read0607(may06); Run;
%Read0607(aug06); Run;
%Read0607(jan07); Run;

Data CPS0307;
  Set CPSfeb03 CPSjun03 CPSnov03 CPSmay06 CPSaug06 CPSjan07;
  If (18<=PEAge<=24) Then AgeGrp=1;
  Else If (25<=PEAge<=44) Then AgeGrp=2;
  Else If (45<=PEAge<=64) Then AgeGrp=3;
  Else If (PEAge>64) Then AgeGrp=4;
  If SmokStat In (2,3) Then Current=1;
  Else If SmokStat In (1,4) Then Current=2;
  If HRYear4=2003 Then SurvGrp=1;
  Else If HRYear4 In (2006,2007) Then SurvGrp=2;
  Format AgeGrp AgeGrpF. PESex PESexF. Current CurrentF.;
  Label AgeGrp   = "Age Group"
        Current  = "Current Cigarette Smoker"
        PESex    = "Gender"
        OccurNum = "Unique Person Identifier"
        QstNum   = "Unique Household Identifier"
        SmokStat = "Smoker Recode"
        SurvGrp  = "Survey Group";
Run;
```

```
Data CPS0307(Drop=I J RepWt001-RepWt160 SmplWgt);
  Set CPS0307;
  Array OldR(160) RepWt001-RepWt160;
  Array NewR(240) NWgt001-NWgt240;
  NSmplWgt=SmplWgt/6;
  If SurvGrp=1 Then Do;
    Do I = 1 to 80;
      NewR(I)=(1/6)*(SmplWgt+(.866025*(OldR(I)-SmplWgt))); /*.866025=1/2x(Sqrt(240/80)*/
      End;
    Do I = 81 to 240;
      NewR(I)=SmplWgt/6;
      End;
    End;
  Else Do;
    Do I = 1 to 80;
      NewR(I)=SmplWgt/6;
      End;
    Do I = 81 to 240;
      J = I - 80;
      NewR(I)=(1/6)*(SmplWgt+(.612372*(OldR(J)-SmplWgt))); /*.612372=1/2x(Sqrt(240/160)*/
      End;
    End;
Run;
```

```
Proc Crosstab Data=CPS0307 Design=BRR;      /* SUDAAN CrossTab Procedure */
  Setenv DecWidth=4;
  Weight NSmplWgt;
  RepWgt NWgt001-NWgt240/ADJFay=16;
  Class Current PESex AgeGrp/Nofreqs;
  Tables PESex*Current AgeGrp*Current;
  RTitle "Tobacco Use Supplement to The Current Population Survey";
  RTitle "February 2003, June 2003, November 2003, May 2006, August 2006 and January 2007";
  RTitle "Interviewed, Adults, Ages 18 and Over";
  RTitle "Self and Proxy Respondents";
  RTitle "Current Cigarette Smoking Status By Gender and Age Group";
  Print NSum = "Sample Size"
        WSum = "Population Size"
        RowPer = "Percent"
        SERow = "Standard Error"
        LowRow = "Lower 95% CI"
        UpRow = "Upper 95% CI"
        /NSumFmt=F8.0 WSumFmt=F10.0;
Run;
```

Variance Estimation Method: BRR
Tobacco Use Supplement to The Current Population Survey
February 2003, June 2003, November 2003, May 2006, August 2006 and January 2007
Interviewed, Adults, Ages 18 and Over
Self and Proxy Respondents
Current Cigarette Smoking Status By Gender and Age Group
by: Gender, Current Cigarette Smoker.

Gender	Current Cigarette Smoker Total	Current Cigarette Smoker Yes	Current Cigarette Smoker No
Total Sample Size	461702	84536	377166
Total Population Size	214976422	38783864	176192558
Total Percent	100.0000	18.0410	81.9590
Total Standard Error	0.0000	0.0911	0.0911
Total Lower 95% CI		17.8622	81.7788
Total Upper 95% CI		18.2212	82.1378
Male Sample Size	216833	44277	172556
Male Population Size	103321845	21047600	82274244
Male Percent	100.0000	20.3709	79.6291
Male Standard Error	0.0000	0.1216	0.1216
Male Lower 95% CI		20.1324	79.3885
Male Upper 95% CI		20.6115	79.8676
Female Sample Size	244869	40259	204610
Female Population Size	111654577	17736264	93918314
Female Percent	100.0000	15.8849	84.1151
Female Standard Error	0.0000	0.0996	0.0996
Female Lower 95% CI		15.6897	83.9178
Female Upper 95% CI		16.0822	84.3103

Tips and Tricks of Handling the TUS DATA Merging Overlap Supplement

- Unique CPS Panel Design Feature
  - Each Household in the Sample is Surveyed for Four Consecutive Months (Panels 1-4) and then for Four Consecutive Months (Panels 5-8) Nine Months Later
  - Persons in Panels 1, 2 or 3 in February 2002 were in Panels 5, 6 or 7 in February 2003

Tips and Tricks of Handling the TUS DATA Merging Overlap Supplement

- Matching Variables for February 2002 and February 2003 Overlap
  - Household Identifier (HRHHID)
  - Month in Sample (HRMIS)
    - 1 in Feb 2002 = 5 in Feb 2003
    - 2 in Feb 2002 = 6 in Feb 2003
    - 3 in Feb 2002 = 7 in Feb 2003
    - No TUS Items for Panels 4 and 8 in Both 2002 and 2003.
  - Sample Identifier (HRSAMPLE)
  - Serial Suffix (HRSERSUF)
  - Household Number (HUHHNum)
  - Persons' Line Number (PULINENO)
  - Gender (PESEX)
  - Persons' Age (PEAGE):
    - Match If 2003 Age is within Plus or Minus 1 of 2002 Age

Tips and Tricks of Handling the TUS DATA Merging Overlap Supplement

- Match Results
  - 22,598: Self and Proxy
  - 15,846: Self Only
- Reasons For Mismatches
  - Migration: Entire Households and Individuals Move to Another Location
  - Individual or Household Non-Response

Tips and Tricks of Handling the TUS DATA
Linking to Other CPS Data/Supplements to the CPS

- TUS-CPS Can Be Linked With Other CPS Basic and CPS Supplement Data.
- Allows the Opportunity to Include Other Topics in Analysis Not Included in The TUS
  - March ASEC, American Time Use Survey (ATUS), Voting and Registration, Computer and Internet Use, Food Security Information
- Example: Starting in January 2003, Occupation and Industry Were Coded Differently than Pre 2003.
  - Questions Were Not Modified
  - Information Gathered Classified According to New Standards and Definitions
  - How to Group New Occupation Coding Scheme into Occupation Groups White Collar, Blue Collar, Service and Other.
  - Merged February 2002 CPS Data With Bureau of Labor Statistics Monthly Extract File for February 2002.

Tips and Tricks of Handling the TUS DATA
Merging January 2007 TUS-CPS with 2007 ASEC

- Example Merging the March 2007 Annual Social and Economic (ASEC) Supplement with the January 2007 Tobacco Use Supplement to The Current Population Survey
- Construct Health Insurance Coverage Status (HIC_Stat) Variable Using ASEC Data. Based on Census Programming Code.
  - http://www.census.gov/hhes/www/hlthins/methodology/programming/cps/recoding.html
  - http://www.census.gov/hhes/www/hlthins/index.html
- Construct Smoking Ban at Work Variable from TUS-CPS Data
- Tables:
  - Health Insurance Coverage Status, ASEC Data
  - Health Insurance Coverage Status, ASEC/TUS-CPS Merged Data
  - Health Insurance Coverage Status X Smoking Ban at Work, Merged Data

Tips and Tricks of Handling the TUS DATA
Merging January 2007 TUS-CPS with 2007 ASEC

- Matching Variables for TUS-CPS and ASEC
  - Month in Sample (HRMIS)
    - 1 in January = 3 in March
    - 2 in January = 4 in March
    - 5 in January = 7 in March
    - 6 in January = 8 in March
  - PERIDNUM in ASEC (Person Identifier)
  - Construct PERIDNUM in TUS-CPS
    - Household Identifier (HRHHID)
    - Household Identifier 2 (HRHHID2)
    - Persons' Line Number (PULINENO)
  - Gender (PESEX)
  - Persons' Age (PEAGE):
    - Match If January Age is within 1 of March Age

```
/*****************************************************************/
/* Program: workshop.example.tuscps.asec.link.sas                */
/*    Date: September 2013                                       */
/*****************************************************************/
Filename CPSJan07 "cpsjan07.dat" lrecl=1313;
Filename ASEC2007 "cpsmar07.dat" lrecl=974;

Proc Format;
  Value HIC_StaF
    1 = "Covered/Insured"
    2 = "Not Covered/Uninsured"
    ;
  Value WorkBanF
    1 = "Ban"
    2 = "No Ban"
    ;
```

```
Data HH_2007(Keep=H_MIS SEQ)
     Per_2007(Drop=H_MIS H_SEQ RecType);
  Infile ASEC2007;  Input @001 RecType   1. @;
  If RecType=1 Then Do;
    Input @002 H_SEQ     5.
          @029 H_MIS     1.;
    Seq=H_SEQ;
    Output HH_2007;
    End;
  Else If RecType=3 Then Do;
    Input @002 PH_SEQ    5.
          @009 A_LineNo  2.
          @015 A_Age     2.
          @020 A_Sex     1.
          @748 HI        1.
          @750 DEPHI     1.
          @757 PRIV      1.
          @759 DEPRIV    1.
          @765 OUT       1.
          @766 CARE      1.
          @767 CAID      1.
          @770 OTH       1.
          @771 (OTYP_1-OTYP_5) (1.)
          @776 OTHSTPER  1.
          @777 (OTHSTYP1-OTHSTYP6) (2.)
          @866 IAHIPER   1.
          @867 (AHITYP1-AHITYP6) (2.)
          @880 PCHIP     1.
          @912 PERIDNUM  $Char22.;
    Seq=PH_SEQ;
    Output Per_2007;
    End;
Run;
```

```
Proc Sort Data=HH_2007;
  By Seq;
Run;

Proc Sort Data=Per_2007;
  By Seq;
Run;

Data ASEC2007;
  Merge HH_2007
        Per_2007(In=IN1);
  By Seq;
  If In1;
Run;

/* Create Health Insurance Variables Based on Census Documentation   */
Data ASEC2007;
  Set ASEC2007;
  If HI=1 | DEPHI=1 | PRIV=1 | DEPRIV=1 | OUT=1 Then COV_HI=1;
  Else If (10<=OTHSTYP1<=14) | (10<=OTHSTYP2<=14) | (10<=OTHSTYP3<=14) |
     (10<=OTHSTYP4<=14) | (10<=OTHSTYP5<=14) | (10<=OTHSTYP6<=14) Then COV_HI=1;
  Else If (10<=AHITYP1<=15) | (10<=AHITYP2<=15) | (10<=AHITYP3<=15) |
    (10<=AHITYP4<=15) | (10<=AHITYP5<=15) | (10<=AHITYP6<=15) Then COV_HI=1;
  Else COV_HI=2;

  If Care=1 Then MCARE=1;
  Else If (OTHSTYP1=1) | (OTHSTYP2=1) | (OTHSTYP3=1) | (OTHSTYP4=1) | (OTHSTYP5=1) |
    (OTHSTYP6=1) Then MCARE=1;
  Else If (AHITYP1=1) | (AHITYP2=1) | (AHITYP3=1) | (AHITYP4=1) | (AHITYP5=1) | (AHITYP6=1)
    Then MCARE=1;
  Else MCARE=2;
```

```
  If CAID=1 | PCHIP=1 Then MCAID=1;
  Else If (OTHSTYP1 In (2,7,9,15) | OTHSTYP2 In (2,7,9,15) | OTHSTYP3 In (2,7,9,15) |
    OTHSTYP4 In (2,7,9,15) | OTHSTYP5 In (2,7,9,15) | OTHSTYP6 In (2,7,9,15)) Then MCAID=1;
  Else If (AHITYP1 In (2,7,9) | AHITYP2 In (2,7,9) | AHITYP3 In (2,7,9) |
     AHITYP4 In (2,7,9) | AHITYP5 In (2,7,9) | AHITYP6 In (2,7,9)) Then MCAID=1;
  Else MCAID=2;

  If OTYP_1=1 | OTYP_2=1 | OTYP_3=1 | OTYP_5=1 Then CHAMP=1;
  Else If (3<=OTHSTYP1<=6) | (3<=OTHSTYP2<=6) | (3<=OTHSTYP3<=6) |
     (3<=OTHSTYP4<=6) | (3<=OTHSTYP5<=6) | (3<=OTHSTYP6<=6) Then CHAMP=1;
  Else If (3<=AHITYP1<=6) | (3<=AHITYP2<=6) | (3<=AHITYP3<=6) |
     (3<=AHITYP4<=6) | (3<=AHITYP5<=6) | (3<=AHITYP6<=6) Then CHAMP=1;
  Else CHAMP=2;

  If COV_HI=1 | MCARE=1 | MCAID=1 | CHAMP=1 Then HIC_Stat=1;
  Else HIC_Stat=2;
  Label COV_HI   = "Private health insurance status"
        MCARE    = "Medicare status"
        MCAID    = "Medicaid status"
        CHAMP    = "Military health care status"
        HIC_Stat = "Health insurance coverage status";
  Format HIC_Stat HIC_StaF.;
Run;

Proc Freq Data=ASEC2007;
  Title1 "2007 Annual Social and Economic (ASEC) Supplement";
  Title2 "Adults Ages 18 to 64";
  Where (18<=A_Age<=64);
  Table HIC_Stat;
Run;

Data ASEC2007;
  Set ASEC2007;
  If H_MIS In (3,4,7,8);
  HRMIS_B=H_MIS;
Run;
```

```
Data CPSJan07;
  Infile CPSJan07;
  Input @0001 HRHHID   $Char15.
        @0016 HRMonth  2.
        @0018 HRYear4  4.
        @0063 HRMIS    2.
        @0071 HRHHID2  $Char5.
        @0122 PEAge    2.
        @0129 PESex    2.
        @0147 PULineNo 2.
        @0161 PRPERTYP 2.
        @0955 PEA1     2.
        @0959 PEA3     2.
        @1240 PEK2A    2.
        @1242 PEK3A    2.
        @1244 PEK3B    2.
        @1278 Intrview 2.
        @1288 PRS64    2.;
  If PRPerTyp=2;     /* Adult Civilian Household Member */
  If Intrview=1;     /* Supplement Interview            */
  If PEAge>17;       /* Ages 18 and over                */
  If HRMIS In (1,2,5,6);
  PERIDNum=HRHHID||HRHHID2||Put(PULineNo,Z2.);
  HRMIS_B=HRMIS+2;
  If PEK2A In (1,2) & PEK3A Not In (-9,-3,-2) & PEK3B Not In (-9,-3,-2) Then Do;
    If PEK3A=1 & PEK3B=1 Then WorkBan=1;
    Else WorkBan=2;
    End;
  Else WorkBan=9;
  Label WorkBan  = "Smoking ban at work";
  Format WorkBan WorkBanF.;
Run;
```

```
Proc Sort Data=CPSJan07;
  By HRMIS_B PERIDNum;
Run;

Proc Sort Data=ASEC2007;
  By HRMIS_B PERIDNum;
Run;

Data AJan2007;
  Merge CPSJan07(In=InTUS)
        ASEC2007(In=InASEC);
  By HRMIS_B PERIDNum;
  If InTUS & InAsec;
  If PESex=A_Sex;
  If (PEAge=A_Age) | (PEAge=(A_Age-1));
Run;

Proc Freq Data=AJan2007;
  Title1 "January 2007 Tobacco Use Supplement to the Current Population Survey Merged With";
  Title2 "2007 Annual Social and Economic (ASEC) Supplement";
  Title3 "Adults Ages 18 to 64";
  Where (18<=PEAge<=64);
  Table HIC_Stat;
Run;

Proc Freq Data=AJan2007;
  Title3 "Adults Ages 18 to 64, Self Respondents";
  Where (18<=PEAge<=64) & WorkBan In (1,2) & (PRS64=1);
  Table HIC_Stat*WorkBan;
Run;
```

2007 Annual Social and Economic (ASEC) Supplement
Adults Ages 18 to 64

Health Insurance Coverage Status

HIC_Stat	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Covered/Insured	100650	80.89	100650	80.89
Not Covered/Uninsured	23776	19.11	124426	100.00

January 2007 Tobacco Use Supplement to the Current Population Survey Merged With 2007 Annual Social and Economic (ASEC) Supplement
Adults Ages 18 to 64

Health insurance coverage status

HIC_Stat	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Covered/Insured	26864	83.80	26864	83.80
Not Covered/Uninsured	5194	16.20	32058	100.00

January 2007 Tobacco Use Supplement to the Current Population Survey Merged With 2007 Annual Social and Economic (ASEC) Supplement
Adults Ages 18 to 64, Self Respondents

Table of HIC_Stat by WorkBan

HIC_Stat	Ban	No Ban	Total
Covered/Insured - Frequency	8679	2235	10914
Covered/Insured - Percent	71.06	18.30	89.36
Covered/Insured - Row Pct	79.52	20.48	
Covered/Insured - Col Pct	90.75	84.37	
Not Covered/Uninsured - Frequency	885	414	1299
Not Covered/Uninsured – Percent	7.25	3.39	10.64
Not Covered/Uninsured - Row Pct	68.13	31.87	
Not Covered/Uninsured - Col Pct	9.25	15.63	
Total - Frequency	9564	2649	12213
Total - Percent	78.31	21.69	100.00