Welcome and thank you for joining. The webinar will begin shortly.
Welcome

<table>
<thead>
<tr>
<th>QUESTIONS &amp; TECHNICAL SUPPORT</th>
<th>Participants are on mute. Use the <strong>Chat Box</strong> to ask questions or request support. Questions will be addressed during designated Q&amp;A periods.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOSED Captioning Service</td>
<td>Closed captioning is available. A link to the service will be provided in the Chat Box.</td>
</tr>
<tr>
<td>Webinar Recording &amp; Materials</td>
<td>The webinar is being recorded. Materials will be posted online in approximately three weeks. Email notification will be sent.</td>
</tr>
</tbody>
</table>
Speaker

James (Todd) Gibson
Senior Programmer/Analyst
Information Management Services (IMS)

gibsont@imsweb.com
301-680-9770
Using SAS to Analyze TUS-CPS Data

James ‘Todd’ Gibson
Information Management Services, Inc

August 5, 2021
Disclaimer

The views and opinions expressed are my own and do not necessarily represent the views, official policy, or position of the U.S. Government, U.S. Department of Health and Human Services or any of its affiliated institutions or agencies.
Agenda

- Obtaining the TUS-CPS data
- Creating SAS datasets for 2018-2019 survey wave data
  - Current cigarette smoking status table
  - Adding replicate weights and calculating standard errors/conf intervals
- Creating the 1992-2019 harmonized data SAS dataset
  - Examples using the harmonized dataset
  - Merging replicate weights to harmonized dataset
- Useful TUS-CPS links/Contact information
- Questions
?? Have a Question ??

- **During the Presentation**
  - Enter any questions related to the current section in the chat box
  - At the end of each section, there will be time (5 minutes) for answering questions related to the current section.

- **At the end of the Presentation**
  - Remaining time will be used to answer questions not covered in the time after each presentation

- **After Webinar**
  - Email question to myself (gibsont@imsweb.com) and Carolyn Reyes-Guzman (Carolyn.reyes-guzman@nih.gov)
Obtaining the TUS-CPS Data
Obtaining the TUS-CPS Data

Data available on the TUS-CPS website: https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires-data
Obtaining the TUS-CPS Data (continued 2)

▪ What’s available on the Questionnaires and Data Page
  ▪ Technical documentation for individual survey waves
  ▪ Users guide for the 2018-2019 data
  ▪ Data files for individual survey waves and the 1992-2019 harmonized file
  ▪ SAS programs to read data files and create SAS datasets
  ▪ Links to the 2010-2019 replicate weights on the Census CPS FTP site
  ▪ Data tables and other reports
Obtaining the TUS-CPS Data for Seminar Examples

- Data needed:
  - Self response replicate weights for the 2018-2019 survey data
    - [https://www.census.gov/data/datasets/time-series/demo/cps/cps-supp_cps-repwgt/cps-tobacco.html](https://www.census.gov/data/datasets/time-series/demo/cps/cps-supp_cps-repwgt/cps-tobacco.html)
  - 1992-2019 harmonized dataset
  - Replicate weights for the harmonized dataset
  - Technical documentation and SAS code.
Creating SAS Datasets for 2018-2019 Survey Wave Data
Creating SAS Datasets for 2018-2019 Survey Wave Data

- Download data and SAS code for 2018-2019 (Done)
- Unzip data files.
- Open SAS program to create SAS datasets for July 2018 and January 2019
- Modify Filename and Libname statements to match where data are stored.
- Run code to create datasets
- Repeat for May 2019 data.
Current Cigarette Smoking Status Table
Current Cigarette Smoking Status Table

- Read July 2018, January 2019 and May 2019 SAS datasets
- Selections
  - Adult Civilian (PrPerTyp=2)
  - TUS Interviews (Intrview=1)
  - Self Respondents (PRS64=1)
- Keep variables needed for analysis
  - Year (HRYear4), Month (HRMonth), Region (GEReg), Age (PrtAge), Hispanic (PEHspNon), Race (PTDTRace), Sex (PESex), Unique household identifiers (QstNum, OccurNum), Smoker Recode (SmokStat), Self Response Weight (PWSRWgt)
Current Cigarette Smoking Status Table (continued)

- Divide self response weight by number of surveys
  - PWSRWgt=PWSRWgt/3

- Construct variables for age group and race/ethnicity
  - AgeGrp: 18-24, 25-44, 45-64, 65 and over
  - RaceEthn: White alone (NH), Black alone (NH), Hispanic, American Indian/Alaska Native alone (NH), Asian/Pacific Islander alone (NH), 2 or more race (NH)

- Generate table of percentages and counts by sex, region, race/ethnicity and age group using proc tabulate
SAS Code For Example 1

```sas
/* Data TUS119; */
Set TUS119;
FRR6Age=FRR6Agt: / "Divide by number of surveys" /
If (t<=FRR6Agt<=64) Then AgeGrp=1; /* Ages 16-24 */
Else If (65<FRR6Agt<=74) Then AgeGrp=2; /* Ages 25-64 */
Else If (75<FRR6Agt<=84) Then AgeGrp=3; /* Ages 45-64 */
Else If (FRR6Agt>84) Then AgeGrp=4; /* Age 65 and over */

If FRR6pEth==3 Then RaceEthn=1; /* Hispanic */
Else If FRR6Race==1 Then RaceEthn=2; /* White alone, Non-Hispanic */
Else If FRR6Race==2 Then RaceEthn=3; /* Black alone, Non-Hispanic */
Else If FRR6Race==3 Then RaceEthn=4; /* American Indian/Alaska Native alone, Non-Hispanic */
Else If FRR6Race==4 Then RaceEthn=5; /* Asian/Pacific Islander alone, Non-Hispanic */
Else RaceEthn=6; /* 2 or more races, Non-Hispanic */

Label AgeGrp = "AgeGrp"
RaceEthn = "Race/Ethnicity";
Form AgeGrp AgeGrpF. GRR6EthGRR6F. FRR6Sex FRR6SexF. RaceEthn RaceEthnF. SmokStat SmokStatF.

/* Data TUS119; */
Set TUS119;
FRR6Sex=FRR6Sex: / "Divide by number of surveys" /
If (FFSex==1) Then Sex=2; /* Female */
Else FFR6Sex=1; /* Male */

If RaceEthn==1 Then RaceEthnF=1; /* White alone, Non-Hispanic */
Else If RaceEthn==2 Then RaceEthnF=2; /* Black alone, Non-Hispanic */
Else If RaceEthn==3 Then RaceEthnF=3; /* American Indian/Alaska Native alone, Non-Hispanic */
Else If RaceEthn==4 Then RaceEthnF=4; /* Asian/Pacific Islander alone, Non-Hispanic */
Else RaceEthnF=5; /* 2 or more races, Non-Hispanic */

Label Sex = "Sex"
RaceEthn = "Race/Ethnicity";
Form FRR6Sex FRR6SexF. RaceEthn RaceEthnF. SmokStat SmokStatF.

/* Data TUS119; */
Set TUS119;
FRR6wht=FRR6wht: / "Divide by number of surveys" /
If (FFR6wht==1) Then Wht=2; /* White alone, Non-Hispanic */
Else FRR6wht=1; /* Black alone, Non-Hispanic */

If FRR6CHW==1 Then CHW=2; /* CHW interview */
Else FRR6CHW=1; /* Self respondents */

Run;
```
### Example 1 Results

#### 2018-19 Tobacco Use Supplement to the Current Population Survey

**Current Cigarette Smoking Status**

<table>
<thead>
<tr>
<th>Current Cigarette Smoking Status</th>
<th>Never</th>
<th>Every day</th>
<th>Some days</th>
<th>Former</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>70.4</td>
<td>8.7</td>
<td>2.7</td>
<td>19.2</td>
<td>248,831,743</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>136,805</td>
</tr>
<tr>
<td>Male</td>
<td>66.8</td>
<td>9.7</td>
<td>3.3</td>
<td>21.2</td>
<td>119,869,783</td>
</tr>
<tr>
<td>Female</td>
<td>74.6</td>
<td>7.8</td>
<td>2.1</td>
<td>15.4</td>
<td>128,971,960</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>69.7</td>
<td>7.8</td>
<td>2.4</td>
<td>20.0</td>
<td>43,592,023</td>
</tr>
<tr>
<td>Midwest</td>
<td>65.2</td>
<td>11.0</td>
<td>2.9</td>
<td>20.8</td>
<td>51,564,592</td>
</tr>
<tr>
<td>South</td>
<td>71.2</td>
<td>9.5</td>
<td>2.8</td>
<td>15.7</td>
<td>94,343,821</td>
</tr>
<tr>
<td>West</td>
<td>74.0</td>
<td>6.2</td>
<td>2.8</td>
<td>17.0</td>
<td>59,331,308</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White (NH)</td>
<td>66.1</td>
<td>10.1</td>
<td>2.5</td>
<td>22.4</td>
<td>156,937,853</td>
</tr>
<tr>
<td>Black (NH)</td>
<td>76.1</td>
<td>8.9</td>
<td>3.7</td>
<td>11.3</td>
<td>29,471,071</td>
</tr>
<tr>
<td>Hispanic</td>
<td>81.7</td>
<td>4.8</td>
<td>2.7</td>
<td>10.8</td>
<td>41,124,089</td>
</tr>
<tr>
<td>American Indian/Alaska Native (NH)</td>
<td>60.2</td>
<td>14.1</td>
<td>6.6</td>
<td>19.0</td>
<td>1,849,561</td>
</tr>
<tr>
<td>Asian/Pacific Islander (NH)</td>
<td>85.0</td>
<td>3.4</td>
<td>2.0</td>
<td>8.8</td>
<td>16,813,772</td>
</tr>
<tr>
<td>2 or more races (NH)</td>
<td>64.2</td>
<td>12.4</td>
<td>4.7</td>
<td>18.7</td>
<td>3,553,387</td>
</tr>
<tr>
<td>Ages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>88.3</td>
<td>4.7</td>
<td>2.7</td>
<td>4.2</td>
<td>29,477,015</td>
</tr>
<tr>
<td>25-44</td>
<td>74.5</td>
<td>9.2</td>
<td>3.4</td>
<td>12.9</td>
<td>85,128,621</td>
</tr>
<tr>
<td>45-64</td>
<td>66.3</td>
<td>11.6</td>
<td>2.7</td>
<td>19.4</td>
<td>82,591,408</td>
</tr>
<tr>
<td>65+</td>
<td>59.8</td>
<td>5.7</td>
<td>1.5</td>
<td>32.9</td>
<td>51,634,620</td>
</tr>
</tbody>
</table>

**Sample Size:** 136,805
Adding Replicate Weights and Calculating Standard Errors and Confidence Intervals
Adding Replicate Weights & Calculating SEs and CIs

- Read replicate weights for July 2018, January 2019 and May 2019
- Concatenate replicate weight datasets
- Divide self response base and replicate weights by number of surveys
- Sort main survey and replicate weights by Year (HRYear4), Month (HRMonth), Unique household identifiers (QstNum, OccurNum)
- Merge main survey and replicate weights by sorted variables
- Check that data merged using proc freq
- Generate percentages, standard errors and 95% confidence intervals using proc surveyfreq
SAS Code For Example 2

```sas
proc sort data=reps1819;
  by HRYear HRMonth QstNum OccurNum;
run;

proc sort data=tus1819;
  by HRYear HRMonth QstNum OccurNum;
run;

data tus1819;
  merge tus1819(in=In1) reps1819(in=In2);
  by HRYear HRMonth QstNum OccurNum;
  if In1=1 and In2=1;
run;

data reps1819;
  set reps1819;
  if SmplWgt=9 then SmplWgt=-9;
run;

data tus1819;
  set tus1819;
  if SmplStat=9 then SmplStat=-9;
run;

/* In the interest of time only doing overall and by sex */
data surveyfreq data=tus1819 varnames=RSS(fay=0.5);
  title "2015-15 Tobacco Use Supplement to the Current Population Survey";
  title2 "Current Cigarette Smoking Status";
  tables SmplStat/CL;
  tables (FESx GESx RaceEthn AgeGrp) SmplStat/CL Row;
  weight SmplWgt;
  reweights RepWt01-RepWt160;
run;
```

---

**Filename** RepJul18 "data\jull18srrep.dat" LRecL=1617;
**Filename** RepJan19 "data\jan19srrep.dat" LRecL=1617;
**Filename** RepMay19 "data\may19srrep.dat" LRecL=1617;

```sas
%macro ReadRep(RepFile,Yr,Mth);
  data &RepFile;
    infile &RepFile;
    input @001 QstNum 5.
    @006 OccurNum 2.
    @008 SmplWgt 10.4
    @016 (RepWt01-RepWt160) (10.4);
    if SmplWgt=-9;
    HRYear=Yr;
    HRMonth=Mth;
%end;
```

```sas
%ReadRep(RepJul18,2018,7); run;
%ReadRep(RepJan15,2019,1); run;
%ReadRep(RepMay19,2019,5); run;
```

---

**NATIONAL CANCER INSTITUTE**

---

**Page 22**
Example 2 Results

2018-19 Tobacco Use Supplement to the Current Population Survey
Current Cigarette Smoking Status

The SURVEYFREQ Procedure

Data Summary

<table>
<thead>
<tr>
<th>Data Summary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations</td>
<td>137471</td>
</tr>
<tr>
<td>Sum of Weights</td>
<td>250038506</td>
</tr>
</tbody>
</table>

Variance Estimation

<table>
<thead>
<tr>
<th>Method</th>
<th>BRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicate Weights</td>
<td>TUS1819R</td>
</tr>
<tr>
<td>Number of Replicates</td>
<td>160</td>
</tr>
<tr>
<td>Fay Coefficient</td>
<td>0.500</td>
</tr>
</tbody>
</table>

Type of smoker recode

<table>
<thead>
<tr>
<th>SMOKSTAT</th>
<th>Frequency</th>
<th>Weighted Frequency</th>
<th>Std Err of Wgt Freq</th>
<th>Percent</th>
<th>Std Err of Percent</th>
<th>95% Confidence Limits for Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>91047</td>
<td>1751336.96</td>
<td>4226.38</td>
<td>70.38%</td>
<td>0.1584</td>
<td>70.04% - 70.72%</td>
</tr>
<tr>
<td>Every day</td>
<td>12954</td>
<td>216925.54</td>
<td>2361.88</td>
<td>8.72%</td>
<td>0.0952</td>
<td>8.52% - 8.90%</td>
</tr>
<tr>
<td>Some days</td>
<td>3616</td>
<td>66988.54</td>
<td>1311.23</td>
<td>2.69%</td>
<td>0.0527</td>
<td>2.58% - 2.79%</td>
</tr>
<tr>
<td>Former</td>
<td>29189</td>
<td>453066.39</td>
<td>3096.60</td>
<td>18.20%</td>
<td>0.1241</td>
<td>17.96% - 18.45%</td>
</tr>
<tr>
<td>Total</td>
<td>136806</td>
<td>2488317.43</td>
<td>6091.60</td>
<td>100.00%</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>

Frequency Missing = 665
## Cigarette Smoking Status - Percentage Estimates (95% Confidence Intervals) for the US Household Population, 18 Years and Older

<table>
<thead>
<tr>
<th></th>
<th>Current % (CI)</th>
<th>Everyday % (CI)</th>
<th>Some Days % (CI)</th>
<th>Former % (CI)</th>
<th>Never % (CI)</th>
<th>Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td>11.4 (11.2-11.6)</td>
<td>8.7 (8.5-8.9)</td>
<td>2.7 (2.6-2.8)</td>
<td>18.2 (18.0-18.5)</td>
<td>70.4 (70.0-70.7)</td>
<td>248,831,743</td>
<td>136,806</td>
</tr>
<tr>
<td><strong>Sex:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td>12.9 (12.6-13.3)</td>
<td>9.7 (9.4-9.9)</td>
<td>3.3 (3.1-3.5)</td>
<td>21.2 (20.9-21.6)</td>
<td>65.8 (65.4-66.3)</td>
<td>119,889,783</td>
<td>62,162</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td>10 (9.7-10.3)</td>
<td>7.8 (7.6-8.1)</td>
<td>2.1 (2.0-2.3)</td>
<td>15.4 (15.1-15.7)</td>
<td>74.6 (74.2-75.0)</td>
<td>128,971,960</td>
<td>74,644</td>
</tr>
<tr>
<td><strong>Region:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Northeast</strong></td>
<td>10.2 (9.7-10.7)</td>
<td>7.8 (7.4-8.3)</td>
<td>2.4 (2.2-2.7)</td>
<td>20 (19.4-20.7)</td>
<td>69.7 (69.1-70.4)</td>
<td>43,592,023</td>
<td>21,794</td>
</tr>
<tr>
<td><strong>Midwest</strong></td>
<td>14 (13.4-14.6)</td>
<td>11 (10.5-11.6)</td>
<td>2.9 (2.7-3.2)</td>
<td>20.8 (20.1-21.5)</td>
<td>65.2 (64.4-66.1)</td>
<td>51,564,592</td>
<td>27,629</td>
</tr>
<tr>
<td><strong>South</strong></td>
<td>12.1 (11.7-12.4)</td>
<td>9.5 (9.2-9.8)</td>
<td>2.6 (2.5-2.8)</td>
<td>16.7 (16.3-17.1)</td>
<td>71.2 (70.6-71.8)</td>
<td>94,343,821</td>
<td>51,146</td>
</tr>
<tr>
<td><strong>West</strong></td>
<td>9 (8.6-9.4)</td>
<td>6.2 (5.8-6.5)</td>
<td>2.8 (2.6-3.1)</td>
<td>17 (16.5-17.5)</td>
<td>74 (73.3-74.7)</td>
<td>59,331,308</td>
<td>36,237</td>
</tr>
</tbody>
</table>

Results Available on TUS-CPS Website (Table 1)

**SAS vs. SUDAAN**

<table>
<thead>
<tr>
<th><strong>SAS: SurveyFreq</strong></th>
<th><strong>SUDAAN: Crosstab</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Proc SurveyFreq Data=TUS1819R VarMethod=BRR(Fay=0.5);</td>
<td>Proc CrosstabData=TUS1819R Design=BRR;</td>
</tr>
<tr>
<td>Title1 &quot;2018-19 TUS-CPS&quot;; Title2 &quot;Current Cigarette Smoking Status&quot;;</td>
<td>Weight SmplWgt;</td>
</tr>
<tr>
<td>Tables PESex*SmokStat/CL Row;</td>
<td>RepWgt RepWt001-RepWt160/ADJFay=4;</td>
</tr>
<tr>
<td>Weight SmplWgt;</td>
<td>Class SmokStat PESex;</td>
</tr>
<tr>
<td>RepWeights RepWt001-RepWt160;</td>
<td>Tables PESex*SmokStat;</td>
</tr>
<tr>
<td></td>
<td>SubPopX SmokStat!=-9;</td>
</tr>
<tr>
<td></td>
<td>Print NSum = &quot;Sample Size&quot;</td>
</tr>
<tr>
<td></td>
<td>WSum = &quot;Weighted Size&quot;</td>
</tr>
<tr>
<td></td>
<td>RowPer = &quot;Percent&quot;</td>
</tr>
<tr>
<td></td>
<td>SERow = &quot;SE Error&quot;</td>
</tr>
<tr>
<td></td>
<td>LowRow = &quot;Lower 95% CI&quot;</td>
</tr>
<tr>
<td></td>
<td>UpRow = &quot;Upper 95% CI&quot;</td>
</tr>
<tr>
<td></td>
<td>/Style=NCHS NSumFmt=F8.0 WSumFmt=F10.0</td>
</tr>
<tr>
<td></td>
<td>RowPerFmt=F6.1 SERowFmt=F5.1</td>
</tr>
<tr>
<td></td>
<td>LowRowFmt=F6.1 UpRowFmt=F6.1;</td>
</tr>
</tbody>
</table>
Creating The 1992-2019 Harmonized Data SAS Dataset
Creating The 1992-2019 Harmonized Data SAS Dataset

- Included with the Data File
  - SAS Programs to create a SAS dataset from the ASCII text file
    - Main: Reads the data and creates the SAS dataset
    - Format: Formats for all variables. Program called by main program
- Other Useful Information
  - Technical Documentation: Overview of CPS, TUS and Harmonized Data
  - Proc Contents of the data file
  - Unweighted frequency tables of all variable
  - Excel table listing the variables and the source variable by survey wave
Creating The 1992-2019 Harmonized Data SAS Dataset (continued)

- Download data and SAS code for the 1992-2019 Harmonized Data (Done)
- Unzip data file and SAS code.
- Open SAS program to create SAS dataset
- Modify Filename and Libname statements to match where data are stored
- Run code to create dataset
Examples Using Harmonized Dataset
Current Cigarette and E-Cigarette Smoking Status

- Read 1992-2019 Harmonized SAS datasets
- Selections: None
  - Only self respondents to the TUS are included in the harmonized file
- Keep variables needed for analysis
  - Survey Wave (SurWave), Year (SurYear), Month (SurMonth), Region (Region), Record ID (RecordID), Sex (Sex), Current Cigarette Smoking Status (CigStat), Current E-cigarette Smoking Status (ECigStat), Self Response Weight (SRWeight)
Current Cigarette and E-Cigarette Smoking Status (continued)

- Divide self response weight by number of surveys in the survey wave
  - 2000 had 2 surveys: SRWeight=PWSRWgt/2
  - All others had 3 surveys: SRWeight=PWSRWgt/3
- Generate table of current **cigarette** smoking with percentages and counts by survey wave and sex using proc tabulate
- Generate table of current **e-cigarette** smoking with percentages and counts by survey wave, sex and region using proc tabulate
Libname MyLib "data\.*;"

%Include "harmonized.tua_cps.1982.through.2015.formats.sas";

Data Harmon;
  Set MyLib.Harmon;
  Keep SurWave SuYear SurMonth Region RecordID Sex CigStat ECigStat SRWeight;
  Run;

  /* The 2000 survey wave has only 2 surveys. All other survey waves have 3 */

Data Harmon;
  Set Harmon;
  If SurWave=6 Then SRWeight=SRWeight/2;
  Else SRWeight=SRWeight/3;
  Run;

Procs Tabulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to the Current Population Survey";
  Title2 "Current Cigarette Smoking Status";
  Where CigStat=9;
  Class SurWave Sex CigStat;
  Var SRWeight;
  Table SurWave=""(All="Total" Sex="")
          CigStat="Current Cigarette Smoking Status"$SRWeight=""Sum"Population"="Convma120
          N="Sample" F=Comma6;
  Run;

Procs Tabulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to the Current Population Survey";
  Title2 "Current E-Cigarette Smoking Status";
  Where ECigStat=9 & SurWave In (9,10);
  Class SurWave Sex Region ECigStat;
  Var SRWeight;
  Table SurWave=""(All="Total" Sex Region)
          ECigStat="Current E-Cigarette Smoking Status"$SRWeight=""Sum"Population"="Convma120
          N="Sample" F=Comma6;
  Run;
### Example 3 Results (partial)

#### Tobacco Use Supplement to the Current Population Survey

**Current Cigarette Smoking Status**

<table>
<thead>
<tr>
<th>Year</th>
<th>Period</th>
<th>Current Cigarette Smoking Status</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1: Never</td>
<td>2: Every day</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Percent</td>
<td>Percent</td>
</tr>
<tr>
<td>1992-1993</td>
<td>Total</td>
<td>52.1</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>1: Male</td>
<td>44.9</td>
<td>22.0</td>
</tr>
<tr>
<td></td>
<td>2: Female</td>
<td>58.6</td>
<td>18.3</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>53.7</td>
<td>19.3</td>
</tr>
<tr>
<td></td>
<td>1: Male</td>
<td>47.0</td>
<td>21.1</td>
</tr>
<tr>
<td></td>
<td>2: Female</td>
<td>59.8</td>
<td>17.7</td>
</tr>
<tr>
<td>1998-1999</td>
<td>Total</td>
<td>55.4</td>
<td>17.9</td>
</tr>
<tr>
<td></td>
<td>1: Male</td>
<td>48.9</td>
<td>19.8</td>
</tr>
<tr>
<td></td>
<td>2: Female</td>
<td>61.3</td>
<td>16.1</td>
</tr>
<tr>
<td>2000</td>
<td>Total</td>
<td>55.8</td>
<td>17.3</td>
</tr>
<tr>
<td></td>
<td>1: Male</td>
<td>49.5</td>
<td>19.1</td>
</tr>
<tr>
<td></td>
<td>2: Female</td>
<td>61.6</td>
<td>15.6</td>
</tr>
<tr>
<td>2001-2002</td>
<td>Total</td>
<td>57.5</td>
<td>16.8</td>
</tr>
<tr>
<td></td>
<td>1: Male</td>
<td>51.6</td>
<td>18.5</td>
</tr>
<tr>
<td></td>
<td>2: Female</td>
<td>62.8</td>
<td>15.1</td>
</tr>
<tr>
<td>2003</td>
<td>Total</td>
<td>61.0</td>
<td>15.2</td>
</tr>
<tr>
<td></td>
<td>1: Male</td>
<td>55.5</td>
<td>17.0</td>
</tr>
</tbody>
</table>
# Example 3 Results

## Tobacco Use Supplement to the Current Population Survey

### Current E-Cigarette Smoking Status

<table>
<thead>
<tr>
<th>Survey wave</th>
<th>Current E-Cigarette Smoking Status</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1: Never</td>
<td>2: Every day</td>
</tr>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
</tr>
<tr>
<td>2014-2015</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>91.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Male</td>
<td>90.5</td>
<td>0.9</td>
</tr>
<tr>
<td>2: Female</td>
<td>92.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Northeast</td>
<td>92.7</td>
<td>0.6</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>89.9</td>
<td>0.9</td>
</tr>
<tr>
<td>3: South</td>
<td>91.7</td>
<td>0.9</td>
</tr>
<tr>
<td>4: West</td>
<td>92.1</td>
<td>0.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2018-2019</th>
<th>Current E-Cigarette Smoking Status</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1: Never</td>
<td>2: Every day</td>
</tr>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
</tr>
<tr>
<td>Total</td>
<td>91.5</td>
<td>1.1</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Male</td>
<td>89.8</td>
<td>1.3</td>
</tr>
<tr>
<td>2: Female</td>
<td>93.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Northeast</td>
<td>92.5</td>
<td>1.1</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>90.0</td>
<td>1.2</td>
</tr>
<tr>
<td>3: South</td>
<td>91.9</td>
<td>1.1</td>
</tr>
<tr>
<td>4: West</td>
<td>91.4</td>
<td>0.9</td>
</tr>
</tbody>
</table>
Merging the Replicate Weights to Harmonized Dataset
Merging Replicate Weights

- Replicate weights for harmonized file are in 3 files
  - 1995-2003: 80 replicate weights
  - 2006-2019: 160 replicate weights

- SAS code available on TUS-CPS website showing how to read the 3 files and merge with the main harmonized SAS dataset
Merging Replicate Weights (continued)

- Read Harmonized SAS dataset
- Read replicate 3 weight files
- Concatenate replicate weight datasets
- Sort harmonized dataset and replicate weights by Year (SurYear), Month (SurMonth), Record ID (RecordID)
- Merge harmonized dataset and replicate weights by sorted variables
- Check that data merged using proc freq
- Example number of cigarettes per day
SAS Code For Example 4

```sas
* Data Reps;
  set Rep9293 Rep9503 Rep9619;
  run;

* Proc Sort Data=Harmon;
  by SurYear SurMonth RecordID;
  run;

* Proc Sort Data=Reps;
  by SurYear SurMonth RecordID;
  run;

* Data Harmon;
  merge Harmon[by=In1]
    reps[by=In2];
  by SurYear SurMonth RecordID;
  merge=In1;
  reps=In2;
  run;

* Proc Freq Data=Harmon;
  title "Check Harmonized File/Replicate Weight Merge";
  table reps/list missing;
  run;

* Data Harmon415;
  set Harmon;
  if SURNAM=9 /'2014-2015 Survey Wave'/
  run;

  /* Divide weights by number of months (surveys) being combined */
  %July 2014, January 2015, May 2015 */

* Data Harmon415;
  set Harmon415;
  array weights[160] RepWe001-RepWe160;
  sumWeights=repWe19/9;
  do i = 1 to 160;
    weights(i)=weights(i)/3;
  end;
  run;
```
SAS Code For Example 4 (continued)

```sas
/* Example by SEX */
/* Proc SurveyMeans Data=harmon1415 VarMethod=BRR (Fay=0.5); */
Var CigFD;
Domain CurrSmk*Sex;
Weight SmprWgt;
RepWeights RepWt001-RepWt160;
Run;
*/
```

```
Data Harmon1415;
  Set Harmon1415;

  If CigStat in (2,3) Then CurrSmk=1; /* Current Cigarette Smoker */
  Else If CigStat in (4,4) Then CurrSmk=0; /* Non-Smoker */
  Else CurrSmk=.;

  If CigStat=2 & (0<=CPFD<89) Then CigPD=CPFD; /* Daily Smokers */
  Else If CigStat=3 & (0<=CPFD<30) Then CigPD=CPFD; /* Non-Daily Smokers */
  Else CigPD=.;

  Label CurrSmk = "Current Cigarette Smoking Status"
  CigPD = "Number of Cigarettes Per Day"
  Format CurrSmk CurrSmkF.
  Run;
```

```sas
/* Proc SurveyMeans Data=harmon1415 VarMethod=BRR (Fay=0.5); */
Var CigPD;
Domain CurrSmk;
Weight SmprWgt;
RepWeights RepWt001-RepWt160;
Run;
```

```sas
/* Example by Sex */
/* */
Proc SurveyMeans Data=harmon1415 VarMethod=BRR (Fay=0.5);*/
Var CigPD;
Domain CurrSmk*Sex;
Weight SmprWgt;
RepWeights RepWt001-RepWt160;
Run;
*/
```
Example 4 Results


The SURVEYMEANS Procedure

Data Summary

<table>
<thead>
<tr>
<th>Number of Observations</th>
<th>163920</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of Weights</td>
<td>241120556</td>
</tr>
</tbody>
</table>

Variance Estimation

<table>
<thead>
<tr>
<th>Method</th>
<th>BRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replicate Weights</td>
<td>HARMON1415</td>
</tr>
<tr>
<td>Number of Replicates</td>
<td>160</td>
</tr>
<tr>
<td>Fay Coefficient</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Error of Mean</th>
<th>95% CL for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>CigPD</td>
<td>Number of Cigarettes Per Day</td>
<td>22536</td>
<td>12.236119</td>
<td>0.072822</td>
<td>12.0923026</td>
</tr>
</tbody>
</table>
Example 4 Results (continued 1)
### Example 4 Results (continued 2)


**The SURVEYMEANS Procedure**

<table>
<thead>
<tr>
<th>CurrSmk</th>
<th>Variable</th>
<th>Label</th>
<th>N</th>
<th>Mean</th>
<th>Std Error of Mean</th>
<th>95% CL for Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Smoker</td>
<td>CigPD</td>
<td>Number of Cigarettes Per Day</td>
<td>5</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current Cigarette Smoker</td>
<td>CigPD</td>
<td>Number of Cigarettes Per Day</td>
<td>25356</td>
<td>12.236119</td>
<td>0.072022</td>
<td>12.0923026-12.379349</td>
</tr>
</tbody>
</table>

#### Domain Analysis by CurrSmk

[Boxplot showing domain analysis by CurrSmk]
Useful TUS-CPS Links
Useful TUS-CPS Links

- **Main Website**

- **Questionnaires and Data Files**

- **FAQ**

- **User Workshops & Webinars**
Contacts
Contacts

- NCI:
  - ncidccpsbrpadvances@mail.nih.gov
  - Carolyn Reyes-Guzman: Carolyn.reyes-guzman@nih.gov

- IMS:
  - Todd Gibson: gibsont@imsweb.com
?? Questions ??
RAISE YOUR HAND if you wish to be unmuted and ask any final questions.

Ensure the Participants Panel is open.

Raise your hand by clicking on the hand icon.

Lower your hand by clicking on the hand icon again.
THANK YOU FOR YOUR PARTICIPATION

WE VALUE YOUR FEEDBACK!

Please share your feedback via a brief survey.
The survey link will be shared via the Chat Box and email.

FOR MORE INFORMATION & HELPFUL RESOURCES

TUS-CPS Website

cancercontrol.cancer.gov/tus-cps

TUS-CPS Email Subscription

cancercontrol.cancer.gov/brp/tcrb/tus-cps#is-newsletter-subscription

TUS-CPS Team Contact

ncidccpsbrpadvances@mail.nih.gov