

National Cancer Institute TUS-CPS Data User Webinar

Using SAS to Analyze TUS-CPS Data

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

Welcome

QUESTIONS & TECHNICAL SUPPORT

Participants are on mute.

Use the **Chat Box** to ask questions or request support.

Questions will be addressed during designated Q&A periods.

CLOSED CAPTIONING SERVICE

Closed captioning is available.

A link to the service will be provided in the Chat Box.

WEBINAR RECORDING & MATERIALS

The webinar is being recorded.

Materials will be posted online in approximately three weeks.

Email notification will be sent.

Speaker



James (Todd) Gibson

Senior Programmer/Analyst

Information Management Services (IMS)

gibsont@imsweb.com

301-680-9770

2021 TUS-CPS Data User Webinar Series

Using SAS to Analyze TUS-CPS Data

James 'Todd' Gibson

Information Management Services, Inc

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

- TUS-CPS website:
<https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/>



The screenshot displays the National Cancer Institute (NIH) website for the Tobacco Use Supplement to the Current Population Survey (TUS-CPS). The page header includes the NIH logo and the text "NATIONAL CANCER INSTITUTE Division of Cancer Control & Population Sciences". A search bar is located in the top right corner. Below the header is a navigation menu with options: "BRP Home", "Funding Opportunities", "Priority Areas", "Research Resources and Tools", "Program Branches", and "About BRP". The main heading reads "The Tobacco Use Supplement to the Current Population Survey". Below this is a breadcrumb trail: "Behavioral Research Program / Tobacco Control Research Branch (TCRB) / TUS CPS". A logo for the "Tobacco Use Supplement Current Population Survey" is shown, featuring stylized figures. The main text describes the TUS-CPS as an NCI-sponsored survey of tobacco use administered as part of the U.S. Census Bureau's Current Population Survey, with data released for the 2018-2019 TUS-CPS (July 2018, January 2019, and May 2019). A blue button labeled "Tobacco Control Funding Opportunities" is visible. On the right side, there is a "News" section with a circular icon containing a calendar and a list of recent updates:

- [UPDATED 1992-2019 TUS-CPS Harmonized Dataset and NEW accompanying replicate weights now available](#)
- [Register for the 2021 TUS-CPS Data User Webinar Series](#)
- New April 2021 publication, [State-Specific Prevalence of Tobacco Product Use Among US Women, Tobacco Use Supplement to the Current Population Survey, 2018-2019](#)
- [New video on navigating the TUS-CPS webpages](#)
- [New 2018-2019 TUS-CPS User Guide for](#)

Obtaining the TUS-CPS Data (continued 1)


- Data available on the TUS-CPS website:
<https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires-data>

TUS-CPS Questionnaires and Data Files

Behavioral Research Program / Tobacco Control Research Branch (TCRB) / TUS CPS / TUS CPS Questionnaires and Data Files

SECTION MENU

- TUS-CPS FAQs
- TUS-CPS Linkages
- TUS-CPS Publication Database
- TUS-CPS Publications
- TUS-CPS Questionnaires and Data Files**
- TUS-CPS Survey Topics
- TUS-CPS Translations
- User Workshops & Webinars

 **Tobacco Use Supplement**
Current Population Survey

2021 TUS-CPS Data User Webinar Series - REGISTRATION OPEN NOW!

Questionnaires

The questionnaires for the 2001-2002, 2003, 2006-2007, 2010-2011, 2014-2015, and 2018-2019 Tobacco Use Supplement to the Current Population Survey (TUS-CPS) are available for download below in two formats. In addition, for the May 2010-2011 special Longitudinal Cohort TUS-CPS, please see below - "2010-11 TUS-CPS" for the May 2010 baseline questionnaire and "May 2011 Follow-up" for the May 2011 follow-up questionnaire.

Data Files

SAS programs for creating a permanent SAS dataset from the .dat and ASCII text files are available for the harmonized and single wave public use files, respectively. For harmonized data files, please refer to the technical documentation Section V. For single wave public use files, please refer to the data table below. Each program reads all variables from the core and the supplement and creates a SAS dataset containing all the variables.

Replicate Weights

The TUS-CPS replicate weights for 1992-2019 are available below. The Census CPS FTP site contains TUS-CPS replicate weight files for 2010-2019, along with CPS files and other supplement files. The CPS questionnaire is available on the CPS Questionnaire page.

2018-2019 TUS-CPS Data Now Available!

Obtaining the TUS-CPS Data (continued 2)

- What's available on the Questionnaires and Data Page
 - TUS-CPS questionnaires for 2001-2002, 2003, 2006-2007, 2010-2011, 2014-2015 and 2018-2019 survey waves
 - 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:
 - 2018-2019 survey data: July 2018, January 2019 and May 2019
 - 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
 - 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

```
LibName CPS "data\.";

Proc Format;
  Value AgeGrpF
    1 = "..18-24"
    2 = "..25-44"
    3 = "..45-64"
    4 = "..65+"
  ;
  Value GERegF
    1 = "..Northeast"
    2 = "..Midwest"
    3 = "..South"
    4 = "..West"
  ;
  Value PEXexF
    1 = "..Male"
    2 = "..Female"
  ;
  Value RaceEthF
    1 = "..White (NH)"
    2 = "..Black (NH)"
    3 = "..Hispanic"
    4 = "..American Indian/Alaska Native (NH)"
    5 = "..Asian/Pacific Islander (NH)"
    6 = "..2 or more races (NH)"
  ;
  Value SmokStaF
    -9 = "Indeterminant"
    1 = "Never"
    2 = "Every day"
    3 = "Some days"
    4 = "Former"
  ;

Data TUS1819;
  Set CPS.CPSJul18
    CPS.CPSJan19
    CPS.CPSMay19;
  If PRPerTyp=2; /* Adult Civilian Records */
  If Interview=1; /* TUS-CPS Interview */
  If PRS64=1; /* Self Respondents */;
  Keep HRYear4 HRMonth GEReg PrtAge PEHspNon PIDTRace PEXex QstNum OccurNum PEA1 PEA3 SmokStat FWSRWgt;
Run;
```

```
Data TUS1819;
  Set TUS1819;
  FWSRWgt=FWSRWgt/3; /* Divide by number of surveys */

  If (18<=PrtAge<=24) Then AgeGrp=1; /* Ages 18-24 */
  Else If (25<=PrtAge<=44) Then AgeGrp=2; /* Ages 25-44 */
  Else If (45<=PrtAge<=64) Then AgeGrp=3; /* Ages 44-64 */
  Else If (PrtAge>=65) Then AgeGrp=4; /* Age 65 and over */

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

  Label AgeGrp = "Ages"
    RaceEthn = "Race/Ethnicity";
  Format AgeGrp AgeGrpF. GEReg GERegF. PEXex PEXexF. RaceEthn RaceEthF. SmokStat SmokStaF.;
Run;

Proc Tabulate Data=TUS1819 Missing;
  Title1 "2018-19 Tobacco Use Supplement to the Current Population Survey";
  Title2 "Current Cigarette Smoking Status";
  Where SmokStat=-9;
  Class PEXex GEReg RaceEthn AgeGrp SmokStat;
  Var FWSRWgt;
  Table All="Total" PEXex GEReg RaceEthn AgeGrp,
    SmokStat="Current Cigarette Smoking Status"*FWSRWgt="**PctSum<SmokStat>="Percent"*F=7.1
    FWSRWgt="**Sum="Population"*F=Comma2.
    N="Sample"*F=Comma5.;
Run;
```

Example 1 Results

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

	Current Cigarette Smoking Status				Population	Sample
	Never	Every day	Some days	Former		
	Percent	Percent	Percent	Percent		
Total	70.4	8.7	2.7	18.2	248,831,743	136,806
Sex						
..Male	65.8	9.7	3.3	21.2	119,859,783	62,162
..Female	74.6	7.8	2.1	15.4	128,971,960	74,644
Region						
..Northeast	69.7	7.8	2.4	20.0	43,592,023	21,794
..Midwest	65.2	11.0	2.9	20.8	51,564,592	27,629
..South	71.2	9.5	2.6	16.7	94,343,821	51,146
..West	74.0	6.2	2.8	17.0	59,331,308	36,237
Race/Ethnicity						
..White (NH)	65.1	10.1	2.5	22.4	156,937,853	99,777
..Black (NH)	76.1	8.9	3.7	11.3	29,471,071	13,024
..Hispanic	81.7	4.8	2.7	10.8	41,124,089	14,846
..American Indian/Alaska Native (NH)	60.2	14.1	6.6	19.0	1,849,561	1,239
..Asian/Pacific Islander (NH)	85.8	3.4	2.0	8.8	15,813,772	6,226
..2 or more races (NH)	64.2	12.4	4.7	18.7	3,635,397	1,694
Ages						
..18-24	88.3	4.7	2.7	4.2	29,477,095	7,616
..25-44	74.5	9.2	3.4	12.9	85,128,621	43,887
..45-64	66.3	11.6	2.7	19.4	82,591,408	46,757
..65+	59.8	5.7	1.5	32.9	51,634,620	38,546

*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

```
Filename RepJul18 "data\jull18srrep.dat" LRecl=1617;
Filename RepJan19 "data\jan19srrep.dat" LRecl=1617;
Filename RepMay19 "data\may19srrep.dat" LRecl=1617;
```

```
%Macro ReadRep(RepFile,Yr,Mth);
  Data &RepFile;
    Infile &RepFile;
    Input @001 QstNum 5.
           @006 OccurNum 2.
           @008 SmplWgt 10.4
           @018 (RepWt001-RepWt160) (10.4);
    If SmplWgt^=0;
    HRYear4=&Yr;
    HRMonth=&Mth;
%Mend ReadRep;
```

```
%ReadRep(RepJul18,2018,7); Run;
%ReadRep(RepJan19,2019,1); Run;
%ReadRep(RepMay19,2019,5); Run;
```

```
Data Repls1819;
  Set RepJul18
      RepJan19
      RepMay19;
Run;
```

```
Data Repls1819;
  Set Repls1819;
  Array Wgts(161) SmplWgt RepWt001-RepWt160;
  Do I = 1 to 161;
    Wgts(I)=Wgts(I)/3;
  End;
  Drop I;
Run;
```

```
Proc Sort Data=Repls1819;
  By HRYear4 HRMonth QstNum OccurNum;
Run;
```

```
Proc Sort Data=TUS1819;
  By HRYear4 HRMonth QstNum OccurNum;
Run;
```

```
Data TUS1819R;
  Merge TUS1819(In=In1)
        Repls1819(In=In2);
  By HRYear4 HRMonth QstNum OccurNum;
  Main=In1;
  Reps=In2;
Run;
```

```
Proc Freq Data=TUS1819R;
  Title "Check Main/Replicate Weight Merge";
  Table Main*Reps/List Missing;
Run;
```

```
Data TUS1819R;
  Set TUS1819R;
  If SmokStat=-9 Then SmokStat=.;
Run;
```

```
/* In the interest of time only doing overall and by sex */
Proc SurveyFreq Data=TUS1819R VarMethod=BRR(Fay=0.5);
  Title1 "2018-19 Tobacco Use Supplement to the Current Population Survey";
  Title2 "Current Cigarette Smoking Status";
  Tables SmokStat/CL;
  * Tables (PESex GEReg RaceEthn AgeGrp)*SmokStat/CL Row;
  Weight SmplWgt;
  RepWeights RepWt001-RepWt160;
Run;
```

Example 2 Results

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

The SURVEYFREQ Procedure

Data Summary	
Number of Observations	137471
Sum of Weights	250038606

Variance Estimation	
Method	BRR
Replicate Weights	TUS1819R
Number of Replicates	160
Fay Coefficient	0.500

Type of smoker recode							
SMOKSTAT	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent	95% Confidence Limits for Percent	
Never	91047	175133696	422638	70.3824	0.1684	70.0499	70.7149
Every day	12954	21692554	236168	8.7178	0.0952	8.5298	8.9058
Some days	3616	6698854	131123	2.6921	0.0527	2.5880	2.7963
Former	29189	45306639	309960	18.2077	0.1241	17.9627	18.4528
Total	136806	248831743	60916	100.0000			
Frequency Missing = 665							

Results Available on TUS-CPS Website (Table 1)

<https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/results/2018-2019/table-1>

Cigarette Smoking Status - Percentage Estimates (95% Confidence Intervals) for the US Household Population, 18 Years and Older^{1,2}

	Current % (CI)	Everyday % (CI)	Some Days % (CI)	Former % (CI)	Never % (CI)	Population	Sample
Total	11.4 11.2-11.6	8.7 8.5- 8.9	2.7 2.6- 2.8	18.2 18.0-18.5	70.4 70.0-70.7	248,831,743	136,806
Sex:							
Male	12.9 12.6-13.3	9.7 9.4- 9.9	3.3 3.1- 3.5	21.2 20.9-21.6	65.8 65.4-66.3	119,859,783	62,162
Female	10 9.7-10.3	7.8 7.6- 8.1	2.1 2.0- 2.3	15.4 15.1-15.7	74.6 74.2-75.0	128,971,960	74,644
Region:							
Northeast	10.2 9.7-10.7	7.8 7.4- 8.3	2.4 2.2- 2.7	20 19.4-20.7	69.7 69.1-70.4	43,592,023	21,794
Midwest	14 13.4-14.6	11 10.5-11.6	2.9 2.7- 3.2	20.8 20.1-21.5	65.2 64.4-66.1	51,564,592	27,629
South	12.1 11.7-12.4	9.5 9.2- 9.8	2.6 2.5- 2.8	16.7 16.3-17.1	71.2 70.6-71.8	94,343,821	51,146
West	9 8.6- 9.4	6.2 5.8- 6.5	2.8 2.6- 3.1	17 16.5-17.5	74 73.3-74.7	59,331,308	36,237

SAS vs. SUDAAN

SAS: SurveyFreq

```
Proc SurveyFreq Data=TUS1819R VarMethod=BRR(Fay=0.5);  
  
Title1 "2018-19 TUS-CPS";  
Title2 "Current Cigarette Smoking Status";  
  
Tables PEXex*SmokStat/CL Row;  
  
Weight SmpIWgt;  
  
RepWeights RepWt001-RepWt160;
```

SUDAAN: Crosstab

```
Proc CrosstabData=TUS1819R Design=BRR;  
  
Weight SmpIWgt;  
  
RepWgt RepWt001-RepWt160/ADJFay=4;  
  
Class SmokStat PEXex;  
  
Tables PEXex*SmokStat;  
  
SubPopX SmokStat!=-9;  
  
Print NSum = "Sample Size"  
      WSum = "Weighted Size"  
      RowPer = "Percent"  
      SERow = "SE Error"  
      LowRow = "Lower 95% CI"  
      UpRow = "Upper 95% CI"  
      /Style=NCHS NSumFmt=F8.0 WSumFmt=F10.0  
      RowPerFmt=F6.1 SERowFmt=F5.1  
      LowRowFmt=F6.1 UpRowFmt=F6.1;
```

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

SAS Code For Example 3

```
Libname MyLib "data\.";

%Include "harmonzd.tus_cps.1992.through.2019.formats.sas";

Data Harmon;
  Set MyLib.Harmon;
  Keep SurWave SurYear 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=4 Then SRWeight=SRWeight/2;
  Else SRWeight=SRWeight/3;
Run;

Proc 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="(PctSum<CigStat>="Percent"*F=7.1
         SRWeight="(Sum="Population"*F=Comma12.
         N="Sample"*F=Comma8.;
Run;

Proc 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="(PctSum<ECigStat>="Percent"*F=7.1
         SRWeight="(Sum="Population"*F=Comma12.
         N="Sample"*F=Comma8.;
Run;
```


Example 3 Results (partial)

**Tobacco Use Supplement to the Current Population Survey
Current Cigarette Smoking Status**

		Current Cigarette Smoking Status				Population	Sample
		1: Never	2: Every day	3: Some days	4: Former		
		Percent	Percent	Percent	Percent		
1992-1993	Total	52.1	20.0	4.4	23.5	186,158,903	227,716
	1: Male	44.9	22.0	4.8	28.3	88,868,339	97,473
	2: Female	58.6	18.3	4.1	19.0	97,290,564	130,243
1995-1996	Total	53.7	19.3	4.2	22.8	191,589,414	186,713
	1: Male	47.0	21.1	4.7	27.2	91,469,055	78,811
	2: Female	59.8	17.7	3.8	18.7	100,120,359	107,902
1998-1999	Total	55.4	17.9	4.2	22.6	197,826,438	175,900
	1: Male	48.9	19.8	4.7	26.6	94,629,672	75,545
	2: Female	61.3	16.1	3.7	18.9	103,196,766	100,355
2000	Total	55.8	17.3	4.5	22.3	200,490,439	124,219
	1: Male	49.5	19.1	5.2	26.1	95,974,204	54,344
	2: Female	61.6	15.6	3.9	18.9	104,516,236	69,875
2001-2002	Total	57.5	16.8	4.2	21.6	203,817,148	184,947
	1: Male	51.6	18.5	4.8	25.0	97,648,152	80,858
	2: Female	62.8	15.1	3.6	18.4	106,168,996	104,089
2003	Total	61.0	15.2	3.7	20.1	211,502,773	183,003
	1: Male	55.5	17.0	4.2	23.3	101,512,619	79,622

Example 3 Results

**Tobacco Use Supplement to the Current Population Survey
Current E-Cigarette Smoking Status**

		Current E-Cigarette Smoking Status				Population	Sample
		1: Never	2: Every day	3: Some days	4: Former		
		Percent	Percent	Percent	Percent		
Survey wave							
2014-2015	Total	91.6	0.8	1.6	6.1	237,807,441	161,694
	Sex						
	1: Male	90.5	0.9	1.7	6.9	114,114,362	72,200
	2: Female	92.6	0.7	1.4	5.3	123,693,080	89,494
	Region						
	1: Northeast	92.7	0.6	1.3	5.4	42,846,390	27,957
	2: Midwest	89.9	0.9	1.8	7.4	50,321,136	35,345
	3: South	91.7	0.9	1.7	5.8	88,799,567	57,872
4: West	92.1	0.7	1.3	5.8	55,840,349	40,520	
2018-2019	Total	91.5	1.1	1.2	6.2	246,463,934	135,564
	Sex						
	1: Male	89.8	1.3	1.5	7.4	118,594,873	61,522
	2: Female	93.0	0.9	1.0	5.1	127,869,062	74,042
	Region						
	1: Northeast	92.5	1.1	1.1	5.4	43,170,607	21,597
	2: Midwest	90.0	1.2	1.3	7.5	51,048,767	27,360
	3: South	91.9	1.1	1.2	5.8	93,335,561	50,621
4: West	91.4	0.9	1.3	6.4	58,908,999	35,986	

*Merging the Replicate Weights to
Harmonized Dataset*

Merging Replicate Weights

- Replicate weights for harmonized file are in 3 files
 - 1992-1993: 48 replicate weights
 - 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

```
Filename Rep9293 "data\harmonzd.tus_cps.1992.through.2019.replicate.wgts.92_93.dat" Lrecl=567;
Filename Rep9503 "data\harmonzd.tus_cps.1992.through.2019.replicate.wgts.95_03.dat" Lrecl=919;
Filename Rep0619 "data\harmonzd.tus_cps.1992.through.2019.replicate.wgts.06_19.dat" Lrecl=1799;
Libname MyLib "data\.";
```

Proc Format;

```
Value CurrSmkF
  0 = "Non-Smoker"
  1 = "Current Cigarette Smoker"
;
```

```
%Include "harmonzd.tus_cps.1992.through.2019.formats.sas";
```

Data Harmon;

```
Set MyLib.Harmon;
Keep SurWave SurYear SurMonth RecordID Sex CigStat CPDD CPDS SRWeight;
Run;
```

Data Rep9293;

```
Infile Rep9293;
Input @001 SurYear 4.
      @005 SurMonth 2.
      @007 RecordID $Char22.
      @029 Smp1Wgt 11.4
      @040 (RepWt001-RepWt048) (11.4);
Run;
```

Data Rep9503;

```
Infile Rep9503;
Input @001 SurYear 4.
      @005 SurMonth 2.
      @007 RecordID $Char22.
      @029 Smp1Wgt 11.4
      @040 (RepWt001-RepWt080) (11.4);
Run;
```

Data Rep0619;

```
Infile Rep0619;
Input @001 SurYear 4.
      @005 SurMonth 2.
      @007 RecordID $Char22.
      @029 Smp1Wgt 11.4
      @040 (RepWt001-RepWt160) (11.4);
Run;
```

Data Reps;

```
Set Rep9293 Rep9503 Rep0619;
Run;
```

Proc Sort Data=Harmon;

```
By SurYear SurMonth RecordID;
Run;
```

Proc Sort Data=Reps;

```
By SurYear SurMonth RecordID;
Run;
```

Data Harmon;

```
Merge Harmon(In=In1)
      Reps(In=In2);
By SurYear SurMonth RecordID;
Harm=In1;
Reps=In2;
Run;
```

Proc Freq Data=Harmon;

```
Title "Check Harmonized File/Replicate Weight Merge";
Table Harm*Reps/List Missing;
Run;
```

Data Harmon1415;

```
Set Harmon;
If SURWAVE=9; /*2014-2015 Survey Wave*/
Run;
```

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

Data Harmon1415;

```
Set Harmon1415;
Array Wgts(160) RepWt001-RepWt160;
Smp1Wgt=Smp1Wgt/3;
Do I = 1 to 160;
  Wgts(I)=Wgts(I)/3;
End;
Run;
```

SAS Code For Example 4 (continued)

```
☐ Data Harmon1415;
  Set Harmon1415;

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

  If CigStat=2 & (0<=CPDD<=99) Then CigPD=CPDD; /* Daily Smokers */
  Else If CigStat=3 & (0<=CPDS<=30) Then CigPD=CPDS; /* Non-Daily Smokers */
  Else CigPD=.;

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

```
☐ Proc SurveyMeans Data=harmon1415 VarMethod=BRR (Fay=0.5);
  Title "Tobacco Use Supplement to the Current Population Survey, 2014-2015";
  Var CigPD;
  Domain CurrSmk;
  Weight SmplWgt;
  RepWeights RepWt001-RepWt160;
Run;
```

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

Example 4 Results

Tobacco Use Supplement to the Current Population Survey, 2014-2015

The SURVEYMEANS Procedure

Data Summary	
Number of Observations	163920
Sum of Weights	241120556

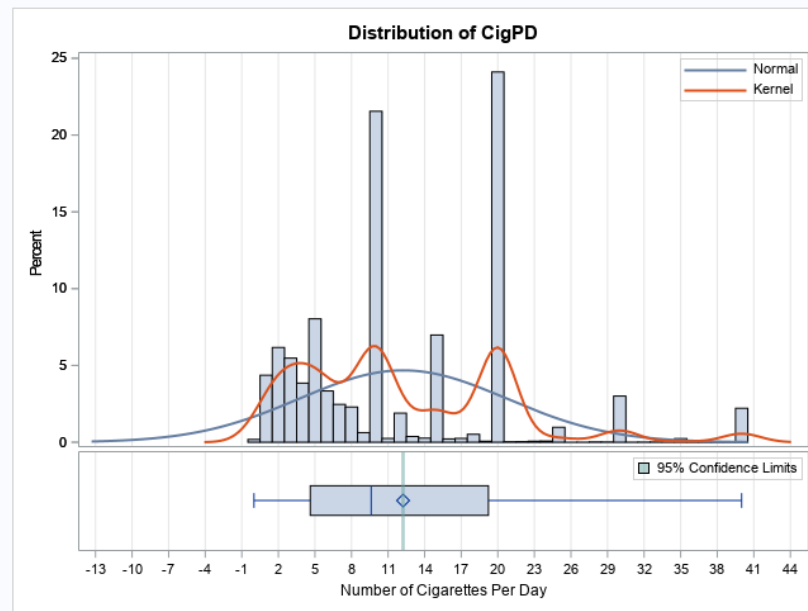
Variance Estimation	
Method	BRR
Replicate Weights	HARMON1415
Number of Replicates	160
Fay Coefficient	0.5

Statistics						
Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
CigPD	Number of Cigarettes Per Day	22536	12.236119	0.072822	12.0923026	12.3799349

Example 4 Results (continued 1)

Tobacco Use Supplement to the Current Population Survey, 2014-2015

The SURVEYMEANS Procedure



Example 4 Results (continued 2)

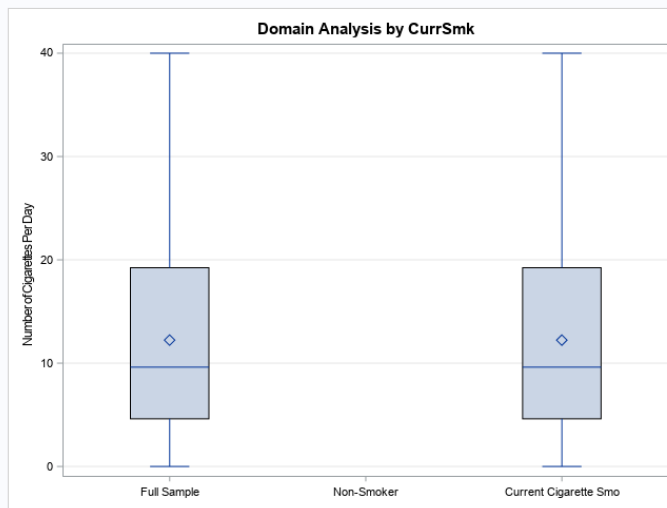
Tobacco Use Supplement to the Current Population Survey, 2014-2015

The SURVEYMEANS Procedure

Statistics for CurrSmk Domains						
CurrSmk	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean
Non-Smoker	CigPD	Number of Cigarettes Per Day	0	.	.	.
Current Cigarette Smoker	CigPD	Number of Cigarettes Per Day	22536	12.236119	0.072822	12.0923026 12.3799349

Tobacco Use Supplement to the Current Population Survey, 2014-2015

The SURVEYMEANS Procedure



Useful TUS-CPS Links

Useful TUS-CPS Links

- Main Website

- <https://cancercontrol.cancer.gov/brp/tcrb/tus-cps>

- Questionnaires and Data Files

- <https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/questionnaires-data>

- FAQ

- <https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/frequently-asked-questions>

- User Workshops & Webinars

- <https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/workshops>

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.

The screenshot shows a Cisco Webex Events window. The main area displays 'Nalini Corcy Host'. On the right, the 'Participants' panel is open, showing a search bar and a list of participants: 'Nalini Corcy Host' and 'Nel C Me'. A hand icon in the bottom right of the Participants panel is highlighted with a red box. At the bottom of the window, the 'Participants' button in the toolbar is also highlighted with a red box. Other buttons include 'Unmute', 'Share', and a red 'X' button.

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



**NATIONAL
CANCER
INSTITUTE**

www.cancer.gov

www.cancer.gov/espanol