Overview & Background of TUS-CPS

Carolyn Reyes-Guzman, PhD, MPH


February 20, 2019
Current Population Survey

- The CPS is a monthly labor force survey conducted in about 56,000 interviewed households across the U.S.
- Each household is interviewed monthly for 4 consecutive months one year, and again for the corresponding time period a year later.
  - Can obtain reliable month-to-month and year-to-year comparisons
- Secondary purpose of CPS is collecting information on demographic characteristics (age, sex, race, marital status, educational attainment, family relationship, occupation, and industry)
  - Additional questions included as Supplements on health, education, income, and previous work experience
TUS-CPS Background

- NCI primarily sponsored TUS-CPS as part of the US Census Bureau’s Current Population Survey every 3-4 years since 1992-93
- FDA currently a co-sponsor since 2014-15
- CDC co-sponsored from 2001-02 through 2006-07
- Most recent publicly released data are from the 2014-2015 TUS-CPS wave; 2018-2019 wave currently being fielded.
  - Data collected during 3 time points for every cycle, except 2000 (2 points)
TUS-CPS Longitudinal Cohort

- NCI additionally sponsored a special longitudinal TUS-CPS cohort administered in both May 2010 and May 2011

- Part of 2001-2002 wave (panels 1-3 of February 2002) overlap with part of 2003 wave (panels 5-7 of February 2003), providing another opportunity to analyze longitudinal data

- Overlap between January and May 1999 and January and May 2000, respectively
What is TUS useful for?

- TUS data can be used by researchers to:
  - Monitor tobacco control progress and assess long-term cross-sectional population trends
  - Track topics such as tobacco health disparities, which may require larger sample sizes
  - Evaluate tobacco control programs
  - Examine national, state and county*-level data

- Data from the TUS can be linked to other CPS supplements

*where available
Recently-Added TUS Topics

- More recent series have included more detailed information on:
  - emerging tobacco products
  - use of flavored non-cigarette tobacco products
  - attitudes towards multi-unit housing smoke-free policies (since 2014-15)
    - These questions not included in current harmonized data file, but will be incorporated after 2018-19 data becomes available
New Questions in 2018-2019 cycle

- Expanded non-cigarette products flavor to get at **flavor categories**, including for e-cigarette “no flavor” vs “tobacco or menthol”

- **Device type** for e-cigarettes

- **E-cigarette Price** in addition to cigarette price (existing)

- Identify **18-34 yr. olds** who **haven’t smoked 100+ cigarettes** but do currently **smoke** cigarettes

- Added back **some treatment aid** items that were previously dropped

- “**Vape-free**” home rules and “**vape-free**” workplace policies

- **Awareness** item for the court-mandated **TI corrective statements** to be placed in newspapers & disseminated through other vehicles.
Purposes for assessing trends within a repeated, cross-sectional tobacco survey

1) Descriptive: Assessing progress in an area of public health
   • Reaching Health People Year 20X0 Objectives

2) Predictive: “Trend Analysis” attempts to project future behavior
   • Predicting when the US will achieve <10% adult cigarette smoking

3) Differential: Showing how a trend over time varies for a particular class or demographic group
   • Reduction in smoking by race/ethnicity

4) Inferential: Relating discontinuities (jumps) to critical events
   • Increase in high school e-cig use after introduction of JUUL
Figure 1. Tobacco Use Supplement to the Current Population Survey: Total U.S. and Regional Trends in Current (Every Day or Some Day) Cigarette Smoking Prevalence, 1992–2015
Figure 3. Trends in Percentage of Current Smokers Reporting Menthol as Usual Type of Cigarette, Total United States, by Sex, TUS-CPS 2003–2015
Figure 4. Trends in Percentage of Current Smokers Reporting Menthol as Usual Type of Cigarette, by Race/Ethnicity, TUS-CPS 2003–2015
Figure 6. Trends in Workplace Smoking Ban, by Region, TUS-CPS 1992–2015

- Total
- Northeast
- Midwest
- South
- West
Figure 7. Home and Workplace Smoking Bans Across TUS-CPS Waves, 1992–2015

Prevalence of smoking bans (%)

TUS-CPS year


43.1 46.1 61.1 67.2 74.2 79.1 83.9 86.5

53.3 63.5 69.0 70.8 77.0 79.1 82.2 79.7
Figure 8. Trends in Attitudes Toward Smoking Bans
In Indoor Workplaces and in Bars/Cocktail Lounges, TUS-CPS 1992–2015
Unique Aspects: Stems from the “3 L’s”

- **LARGE** sample SIZE, **Longevity**, and **Linkages** are KEY to many of the **Unique** features along with **Unique Content** and **Design**
  - Derived from the CPS content
  - Examine detailed tobacco health disparities
  - Study by **small / novel** units of geography
  - CPS design provides opportunity for longitudinal/cohort evaluation
  - CPS **Linkages** through NLMS & other CPS supplements
- **Unique** features’ applications often combine > 1 of these 5 aspects
Uniqueness: Derived from the CPS Content

- Deep Dive on Occupation & Industry
  - bartenders, construction workers, nurses subcategories (LPNs, RNs, etc.)
  - food service workers by job category & type of venue (food prep, dishwashers, waitresses/waiters, servers, hosts/ restaurants, cafeterias)

- Work Force/ Employment Considerations
  - Congressional Budget Office used TUS to assess impact of raising the Federal tobacco excise tax - effect of smoking on earnings perspective
  - use of “# of sick days” as a health outcome in tobacco control studies
  - immigration and access to fringe benefits by Maclean et. al., 2018
Uniqueness: Detailed Tobacco Health Disparities

- Racial/ethnic differences
  - Hispanics by country of origin and geography (e.g., Hispanics in NYC)
  - South Asians (India, Pakistan, Sri Lanka, Bangladesh)
  - Intergenerational differences between blacks by country of origin
- Rural/urban differences
  - 10 HHS regions across the U.S.
  - Restricted data to study rural MS Delta & Appalachia with a new isolation index
- Immigration & acculturation
- Other socio-economic differences
  - Projected impacts of cigarette floor price & excise tax policies (Golden et. al., 2016)
Uniqueness: Small/novel Geographic units

- State and sub-state levels, e.g., major cities are in the sample with certainty
- Major media markets identified which may cross state boundaries
  - Emery et. al., 2012 linked smoking related television advertising ratings to TUS data
  - Niederdeppe et. al., 2018 market level exposure to state anti-smoking campaigns and public support for tobacco control policy
- Comparison of states/ groups of states by tobacco control policy, attitudes, social norms impact on tobacco use/cessation treatment
  - Pierce et. al., 2018- changes in tobacco control policy, norms in CA vs. rest of US- impact on long-term trends in per capita cigarette consumption
  - Dahne et. al., 2017 –studied state tobacco control policies differential impact on cessation treatment utilization across established tobacco disparities groups
Uniqueness: Longitudinal/cohort evaluation from CPS design

- CPS panel design allowed us to efficiently conduct 3 short-term cohorts’ evaluations
  - Vijayaraghavan et. al., 2018 – studied income disparities and diffusion of smoke-free homes impact on smoking cessation using two cohorts
  - Chang et. al., 2018, Zhu et. al., 2009– transitions between cigarette and smoke-less tobacco use from 2002-03 and 2010-11, and 2002-03 cohort data, respectively
  - Leas et. al., 2018 – effectiveness of pharmaceutical smoking cessation aids in a cohort of smokers 2010-11 and 2002-03
  - Nagelhout et. al., 2015 – smokers support smoke-free laws to help them quit smoking, 1999-2000 cohort analysis
- Reliability of self-reported smoking history measures
  - Soulakova et. al., 2012-2018
Uniqueness: CPS Linkages through CPS Supplements

- Special direct linkages to disease outcomes data with National Longitudinal Mortality Study (NLMS)
- Linkages to the CPS Annual Social and Economic Supplement (March ASEC) through the CPS panel design
- CPS data linked to National Death Index (through 2011), NCI SEER Cancer Registry, and CMS Medicare
Uniqueness: CPS Linkages through CPS Supplements

- Linkage to ASEC providing detailed economic, occupational, social, and health insurance data
  - examine effect of Medicaid coverage of tobacco dependence treatments on quitting attempts and intention to quit by Medicaid smokers (Liu et. al, 2010)
  - attitudes towards multi-unit housing smoke-free policies, tobacco use and type of housing
- Linkage to the American Time Use Survey (ATUS) sponsored by BLS
  - provides time estimates that Americans spend in various primary activities for a given day of the week based on a 24 hr. activity diary
  - sample drawn from CPS samples 1 month after they have completed their 8 mo. panel participation
  - Song, 2012- pattern of daily activities for smokers, former smokers, and never smokers
Uniqueness: CPS Linkages through CPS Supplements

- **Voting** and Registration can provide data about the ability to influence policy
  - Hersch et. al., 2004 – compare influence of TUS attitudes to smoking in public places with state smoke-free laws and participation in voting

- **Food Insecurity**
  - Farrelly et. al., 2017 – Trends in food insecurity and cigarette smoking 1998-2011

- **Veterans** – info about branch, length, and period of military service

- **Internet and Computer Use**
  - R21 (Stillman & Wewers described earlier) Supplement (Stillman & Emery)- merging info from this to TUS to develop feasible future interventions in the MS Delta & Appalachia rural/urban areas
Linkages to CPS Supplements

- Voting and Registration: https://www.census.gov/topics/public-sector/voting/data.html
- Computer and Internet Use: https://www.census.gov/topics/population/computer-internet.html
Small Area Estimation
RATIONALE: Why? & What?

- TUS-CPS is designed to produce reliable estimates at national, state, & some sub-state-levels

- Policy makers, cancer control planners and researchers often need county level data for tobacco measures

- **County level** estimates (n=3,137) for the following key measures (2010/2011 TUS, age 18+):
  - % of pop. currently smoking
  - % of pop. that have ever smoked
  - % of pop. quit for 24+ hours within the past year, among those currently smoking at some time during the past year
  - % of pop. reporting a smoke-free workplace policy
  - % of pop. reporting a smoke-free home rule
Small Area Estimates: How?

- Collaboration among NCI, Census Bureau and U. MD

- Use Model-Based Methods, as TUS-CPS has some counties with small (rep.< 100,000 pop.) or zero sample size

- **Borrowing strength** from relevant sources (Census/ Admin. info, related surveys)

- Borrowed strength from covariates & other counties with similar features

- The pool of auxiliary variables include:
  - 30 **county-level** demographic & socio-economic variables from ACS 2005-12, Census 2000 & 2010, & other admin. records
  - 5 **state level** tobacco policy data (cigarette taxes, clean air laws, tobacco control funding, Medicaid coverage for tobacco-related treatment, year that Quitline service initiated)
APPLICATIONS: Current Smoking Map

Current Smoking Prevalence (Age 18+)
US by County, 2010–2011

NCI, Small Area Estimates, TUS/CPS
APPLICATIONS: Home Rule Map
APPLICATIONS: Smoke-free Workplace Policy Maps - 1st TUS Modeled Reports & 2nd TUS State Reports from State Cancer Profiles*

*(statecancerprofiles.cancer.gov/map/map.noimage.php)*
MORE INFO: TUS SAE Data Accessibility & Contact

  - County level current & ever smoking prevalence estimates available upon request *
- NCI SAE general website [sae.cancer.gov](http://sae.cancer.gov)

- **Future**: 2014-15 TUS wave; county estimates by race/ethnicity
- **Contact** Benmei Liu: [liub2@mail.nih.gov](mailto:liub2@mail.nih.gov); our website-SAE details PPT

* avoid confusion as similar estimates for these 2 items derived from combining NHIS & BRFSS released at [https://sae.cancer.gov/nhis-brfss/](https://sae.cancer.gov/nhis-brfss/)
TUS-CPS Unique Characteristics & Linkages
Harmonization Process

Carolyn Reyes-Guzman
NCI, DCCPS

James ‘Todd’ Gibson
Information Management Services, Inc

February 20, 2019
Effort Behind Data Harmonization

- Researchers who previously wanted to examine long-term data had to track different variable names over several survey waves for a specific question of interest.

- Questions on special topics with smaller sample sizes were challenging to study, creating an opportunity for using harmonized data to pool samples over time.

- Data harmonization involved creating one variable name to capture all survey waves, using a “flag” variable to track the survey year.

- “One-stop shopping:” harmonized public use file will be posted on the TUS-CPS website along with supporting documentation (data dictionary, dataset contents, descriptive tables and technical documentation).
Overview of TUS-CPS Harmonized Data File

- **Variables:**
  - **Current Population Survey (CPS) Core:**
    - Region, state, county (limited), family income, metropolitan status, age, sex, race*, Hispanic origin, marital status, education, employment status, nativity (respondent, mother, father)
  - **Tobacco Use Supplement (TUS):**
    - Cigarette Use – including menthol since 2003
    - Workplace and home smoking restrictions
    - Attitudes towards smoke-free policies in indoor work areas/public places
    - Advice by physician/dentist
    - Health perceptions/beliefs (harm reduction)
    - Smoking history, cessation, former smokers
    - Other tobacco products: cigar, pipe and smokeless tobacco use
Data harmonization processes

- Criteria for inclusion: question asked in >1 survey*
- Harmonized file includes adult (18+) self-reported responses, no proxy responses, as they bring greater reliability issues
- Special variable considerations (e.g., when item wording changed over time or when the universe of respondents slightly varied over time)
- Harmonized variable names capture a single question over multiple years
- Variable crosswalk provides detailed notes for the data user on rationale for harmonized variables for special considerations

*few variables were excluded due to substantial question wording changes or difficulty harmonizing
<table>
<thead>
<tr>
<th>Harmonized Codebook Screenshot</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Survey wave</td>
<td></td>
<td>SUMWAVE</td>
<td>SUMWAVE</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
</tr>
<tr>
<td>Month of interview</td>
<td></td>
<td>SUMWAVE</td>
<td>SUMWAVE</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
</tr>
<tr>
<td>Year of interview</td>
<td></td>
<td>SUMWAVE</td>
<td>SUMWAVE</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
<td>H1MENOTH</td>
</tr>
<tr>
<td>Family income</td>
<td>FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
<td>H1FAMINC</td>
</tr>
<tr>
<td>Number of persons in household</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td>NUMHUM</td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td>REG</td>
<td>REG</td>
<td>REG</td>
<td>REG</td>
<td>REG</td>
<td>REG</td>
<td>REG</td>
<td>REG</td>
<td>REG</td>
<td>Volume was constructed. This code must be used with a state code (FIPS) in order to uniquely identify a county. Not all counties identified. Constructed variable in harmonized data file is STCOUNTY. This variable includes state and county.</td>
</tr>
<tr>
<td>State name</td>
<td>STATE</td>
<td>GEOCO</td>
<td>GEOCO</td>
<td>GEOCO</td>
<td>GEOCO</td>
<td>GEOCO</td>
<td>OTCO</td>
<td>OTCO</td>
<td>OTCO</td>
<td></td>
</tr>
<tr>
<td>Metropolitan status</td>
<td>METSTAT</td>
<td>H1METSTAT</td>
<td>H1METSTAT</td>
<td>H1METSTAT</td>
<td>H1METSTAT</td>
<td>H1METSTAT</td>
<td>H1METSTAT</td>
<td>H1METSTAT</td>
<td>H1METSTAT</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>AGE</td>
<td>AGE</td>
<td>AGE</td>
<td>AGE</td>
<td>AGE</td>
<td>AGE</td>
<td>AGE</td>
<td>AGE</td>
<td>AGE</td>
<td>Slight changes in response categories over time.</td>
</tr>
<tr>
<td>Sex</td>
<td>SEX</td>
<td>SEX</td>
<td>SEX</td>
<td>SEX</td>
<td>SEX</td>
<td>SEX</td>
<td>SEX</td>
<td>SEX</td>
<td>SEX</td>
<td></td>
</tr>
<tr>
<td>Race (1992-2002)</td>
<td>RACE2</td>
<td>RACE2</td>
<td>RACE2</td>
<td>RACE2</td>
<td>RACE2</td>
<td>RACE2</td>
<td>RACE2</td>
<td>RACE2</td>
<td>RACE2</td>
<td>Only 4-5 response categories were available to respondents.</td>
</tr>
<tr>
<td>Race (2003-2015)</td>
<td>RACE3</td>
<td>RACE3</td>
<td>RACE3</td>
<td>RACE3</td>
<td>RACE3</td>
<td>RACE3</td>
<td>RACE3</td>
<td>RACE3</td>
<td>RACE3</td>
<td>Only 4-5 response categories were available to respondents.</td>
</tr>
<tr>
<td>Hispanic or non-Hispanic origin</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>HISPAN</td>
<td>1992-1998 asked specificity of origin; 1998- asked as dichotomous variable.</td>
</tr>
<tr>
<td>Highest Level of School Completed or Degree Received</td>
<td>EDU</td>
<td>EDU</td>
<td>EDU</td>
<td>EDU</td>
<td>EDU</td>
<td>EDU</td>
<td>EDU</td>
<td>EDU</td>
<td>EDU</td>
<td>Equivocation was not asked in 2003 but variable was constructed.</td>
</tr>
<tr>
<td>Country of Birth</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td>NATIVITY</td>
<td></td>
</tr>
<tr>
<td>Mother's Country of Birth</td>
<td>MBCR</td>
<td>MBCR</td>
<td>MBCR</td>
<td>MBCR</td>
<td>MBCR</td>
<td>MBCR</td>
<td>MBCR</td>
<td>MBCR</td>
<td>MBCR</td>
<td></td>
</tr>
<tr>
<td>Father's Country of Birth</td>
<td>FBCR</td>
<td>FBCR</td>
<td>FBCR</td>
<td>FBCR</td>
<td>FBCR</td>
<td>FBCR</td>
<td>FBCR</td>
<td>FBCR</td>
<td>FBCR</td>
<td></td>
</tr>
<tr>
<td>Citizenship Status</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>CITIZEN</td>
<td>Slight changes in response categories over time.</td>
</tr>
</tbody>
</table>
Examples of Decision-Making Scenarios

- Former smokers section: the universe of respondents slightly changed over time; some years asked “in the last 3 years”, whereas later years asked “in the last 5 years.”
  - To harmonize, the length of quit (LOQ) was used to set the universe to include only the last 3 years
- Length of quit items: LOQ in earlier surveys was truncated at 12 months, but not truncated in later years.
  - To harmonize, later survey years were truncated to 12 months
Examples of Decision-Making Scenarios

- Quit attempts section: earlier years first asked “Have you ever made a serious attempt to stop smoking…” then “During the past 12 months, have you made a serious attempt to stop smoking…”, but later years asked in the reverse.
  - It isn’t clear whether the order could have impacted patterns of responses.
  - To harmonize, the reverse order was imputed and a note added to the data dictionary to clarify.
Examples of Decision-Making Scenarios

- Ever and current use of any other tobacco products (e-cigarette items were asked starting in 2014)
  - Without more years of data, the variable is not as useful so these items were thus dropped
- Time to first use of a non-cigarette tobacco product: a summary variable was deemed too difficult to harmonize.
  - This will be a follow-up item once the 2018-19 data are available for possible harmonization then.
Overview of Tobacco Use Supplement (TUS) to the Current Population Survey (CPS) Harmonized Data File
Overview of TUS-CPS Harmonized Data File

- Data from the 9 survey waves:
     - January and May 2000 survey was done because of change in wording in 1998-99. Limited number of questions asked. Additional information available in the 2000 technical use documentation.
  8. 2010-2011: May 2010, August 2010 and January 2011
Overview of TUS-CPS Harmonized Data File

- January and May 2000
  - The September 1998, January 2000 and May 2000 supplements have different wording for the "other tobacco use" questions than the 1992-93, 1995-96 and January and May 1999 supplements.
  - In September 1998, the questions were changed to more appropriately capture some day/occasional usage of other tobacco products by changing the previous wording:
    
    From: (Have/Has) (you/name) EVER used pipes, cigars, chewing tobacco or snuff on a regular basis?
    To: (Have/Has) (you/name) EVER used pipes, cigars, chewing tobacco or snuff?
  - In January and May 1999 the wording reverted back to "(Have/Has) (you/name) EVER used pipes, cigars, chewing tobacco or snuff on a regular basis?"
  - It is recommended that for more accurate estimates of other tobacco products the January 2000 and May 2000 data should be averaged or January 2000 and May 2000 be averaged with September 1998 other tobacco product data.
Overview of TUS-CPS Harmonized Data File

From TUS-CPS Website:

**Race/Ethnicity changes in 2003 CPS:** In 2003, significant changes were made to race/ethnicity questions in the CPS. Respondents were able to select more than one race when answering the survey. This change in wording does not impact smoking estimates and trends made for the entire nation from the TUSCS-CPS, but it could potentially impact smoking estimates and trends made by race/ethnicity. NCI has developed a method to construct single race-estimates using data from the post-2003 TUSCS-CPS. The method is useful when trends over time are being examined for single race groups using both pre-2003 and post-2003 data. More information is available in the [Bridging Estimates by Race report](https://cancercontrol.cancer.gov/brp/tcrb/tus-cps/race_bridging_071307.pdf), (PDF) which describes the method and gives an initial assessment of the usefulness of the race adjustment. Also, see [Trends in Smoking Prevalence by Race based on the Tobacco Use Supplement to the Current Population Survey](https://pdfs.semanticscholar.org/f538/2ae7f5f41a9708d12fce3530f87360088b6b.pdf) (PDF) for an application of this race bridging.
# Overview of TUS-CPS Harmonized Data File

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 = White</td>
<td>1 = White only</td>
</tr>
<tr>
<td>2 = Black</td>
<td>2 = Black only</td>
</tr>
<tr>
<td>3 = American Indian, Aleut, Eskimo</td>
<td>3 = American Indian, Alaskan Native only</td>
</tr>
<tr>
<td>4 = Asian or Pacific Islander</td>
<td>4 = Asian only</td>
</tr>
<tr>
<td>5 = Other (1992-93 only, 0.3% of sample)</td>
<td>5 = Hawaiian/Pacific Islander only</td>
</tr>
<tr>
<td></td>
<td>6 = White and Black</td>
</tr>
<tr>
<td></td>
<td>7 = White and American Indian, AK Native</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19 = W-B-AI-A</td>
</tr>
<tr>
<td></td>
<td>20 = Other 2 or 3 races</td>
</tr>
<tr>
<td></td>
<td>21 = Other 4 races or all 5 races</td>
</tr>
</tbody>
</table>
# Overview of TUS-CPS Harmonized Data File

<table>
<thead>
<tr>
<th>Tobacco Use Supplement to The Current Population Survey</th>
<th>Harmonized Data File</th>
<th>Variable: RACE03</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sample</td>
<td>Percent</td>
<td>Sample</td>
<td>Percent</td>
</tr>
<tr>
<td>White only</td>
<td>157,192</td>
<td>85.52</td>
<td>146,230</td>
<td>85.01</td>
</tr>
<tr>
<td>Black only</td>
<td>16,148</td>
<td>8.79</td>
<td>15,178</td>
<td>8.82</td>
</tr>
<tr>
<td>American Indian, Alaskan Native only</td>
<td>1,877</td>
<td>1.02</td>
<td>1,710</td>
<td>0.99</td>
</tr>
<tr>
<td>Asian only</td>
<td>5,379</td>
<td>2.93</td>
<td>5,583</td>
<td>3.25</td>
</tr>
<tr>
<td>Hawaiian/Pacific Islander only</td>
<td>448</td>
<td>0.24</td>
<td>496</td>
<td>0.29</td>
</tr>
<tr>
<td>White and American Indian, Alaskan Native</td>
<td>1,700</td>
<td>0.92</td>
<td>1,591</td>
<td>0.92</td>
</tr>
<tr>
<td>Other 2 or more races</td>
<td>1,066</td>
<td>0.58</td>
<td>1,235</td>
<td>0.72</td>
</tr>
<tr>
<td>Total</td>
<td>183,810</td>
<td>100.00</td>
<td>172,023</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Overview of TUS-CPS Harmonized Data File

- Self respondents, Ages 18 and over
- Weights:
  - Main self-response weight included
    - Controlled for age, race, sex, and Hispanic origin and individual state.
    - Weights can be used to generated estimates for each or the 26 surveys (Sept 1992, Jan 1993,…,July 2015) individually, or for
    - Most reliable estimates, especially for states or sub-state level, will come from combining the data from the 2 surveys in 2000 and all 3 collection periods for all other survey waves. When combining data the weight needs to be divided by number of surveys being combined.
  - Replicate weight will be released in the future as a separate file. Code and information on linking with main harmonized data file and use will be supplied with the file.
Overview of TUS-CPS Harmonized Data File

- File specifics:
  - ASCII Text file
  - Number of records: 1,593,413
  - Record Length: 554
  - Number of variables: 225 (199 TUS)
Heterogeneity in tobacco use behaviors among US black populations

Kelvin Choi, PhD, MPH
Launick Saint-Fort, BS
Division of Intramural Research, NIMHD

Wednesday, February 20, 2019
8:30 am – 11:30 am PST
Disclaimer

The opinions and comments expressed during this presentation are the presenter’s own and do not necessarily represent those of the National Institute on Minority Health and Health Disparities, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, and the U.S. Government.
Background

- Smoking remains the primary cause of death of African-Americans in the U.S.
- 3.8 million foreign-born blacks are currently living in the U.S.
  - 8.7% of the U.S. black population in 2013 and 16.5% by 2060
- Previous studies found differences in prevalence of smoking by global region of origin and age of immigration
Objectives

• Explore whether there is heterogeneity in tobacco-use behaviors (including current cigarette and cigar smoking, time to first cigarette, and menthol cigarette use) among U.S. blacks by global region of origin and age at immigration.
Problem!

US population = 100 people
Blacks = 20 people
Foreign-born blacks < 2 people!

Country of birth!
Available in NHIS restricted-use file
TUS-CPS to the rescue!

- Large sample sizes
  - >100,000 per data collection wave
- Country of birth
  - Public use file
- Year of entry to the US
  - Calculate age of immigration
Methods


• Sample (n=47,857):
  • Self-identified blacks (regardless of Latino origin)
  • 18 years or older
  • Self-respondents
  • Global region of birth
    • U.S. (including Guam, Puerto Rico, the U.S. Virgin Islands, and other U.S. island areas): N=43,560
    • Africa (including both North and Sub-Saharan Africa): N=1,911
    • West Indies: N=2,194
    • Europe: N=192
Country of birth

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Code</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>057</td>
<td>United States</td>
<td>060</td>
<td>American Samoa</td>
</tr>
<tr>
<td>066</td>
<td>Guam</td>
<td>158</td>
<td>Armenia</td>
</tr>
<tr>
<td>069</td>
<td>Northern Marianas</td>
<td>159</td>
<td>Azerbaijan</td>
</tr>
<tr>
<td>073</td>
<td>Puerto Rico</td>
<td>160</td>
<td>Belarus</td>
</tr>
<tr>
<td>078</td>
<td>U. S. Virgin Islands</td>
<td>161</td>
<td>Georgia</td>
</tr>
<tr>
<td>100</td>
<td>Albania</td>
<td>162</td>
<td>Moldova</td>
</tr>
<tr>
<td>102</td>
<td>Austria</td>
<td>163</td>
<td>Russia</td>
</tr>
<tr>
<td>103</td>
<td>Belgium</td>
<td>164</td>
<td>Ukraine</td>
</tr>
<tr>
<td>104</td>
<td>Bulgaria</td>
<td>165</td>
<td>USSR</td>
</tr>
<tr>
<td>105</td>
<td>Czechoslovakia</td>
<td>166</td>
<td>Europe, not specified</td>
</tr>
<tr>
<td>106</td>
<td>Denmark</td>
<td>168</td>
<td>Montenegro</td>
</tr>
<tr>
<td>108</td>
<td>Finland</td>
<td>200</td>
<td>Afghanistan</td>
</tr>
<tr>
<td>109</td>
<td>France</td>
<td>202</td>
<td>Bangladesh</td>
</tr>
<tr>
<td>110</td>
<td>Germany</td>
<td>203</td>
<td>Bhutan</td>
</tr>
<tr>
<td></td>
<td></td>
<td>205</td>
<td>Myanmar (Burma)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>206</td>
<td>Cambodia</td>
</tr>
</tbody>
</table>
Methods

• Measures:
  • Age at entry = Year of survey – Age – Year of entry
  • Current tobacco use = some days + every day
    • Cigarettes
    • Cigars
  • Established smokers (100 cigarettes in the lifetime)
    • Former smoker
    • Started regular smoking before 18 years old: Age of regular smoking
  • Current smokers (Some days + every day cigarette smoking)
    • First cigarette within 30 minutes after waking
    • Regular menthol cigarette smoking
Methods

• Measures (continued)
  • Demographics: Survey year, age, gender, educational attainment, income, and US census region.

• Analysis
  • Weighted multivariable logistic regression
Results: Full sample

Global region of origin

Cigarette

US: 17%  Europe: 18%  Africa: 5%  West Indies: 5%

Age at entry

Cigarette

<13 yrs: 9%  >=13 yrs: 5%

Global region of origin

Cigar

US: 3%  Europe: 3%  Africa: 0%  West Indies: 1%

Age at entry

Cigar

<13 yrs: 1%  >=13 yrs: 1%

(Reference: US-born; *=p<0.05)
Results: Established smokers

Global region of origin

- Former smoker:
  - US: 41%
  - Europe: 33%
  - Africa: 51%
  - West Indies: 55%

- Started regular smoking as minor:
  - US: 46%
  - Europe: 42%
  - Africa: 33%
  - West Indies: 45%

Age at entry

- Former smoker:
  - <13 yrs: 35%
  - >=13 yrs: 56%

- Started regular smoking as minor:
  - <13 yrs: 50%
  - >=13 yrs: 36%

(Reference: US-born; *=p<0.05)
Results: Current smokers

(Reference: US-born; *=p<0.05)
Conclusions

• TUS-CPS harmonized data provide great opportunities to study racial/ethnic disparities in tobacco use!

• Heterogeneity by global region of origin
  • Not all blacks are the same
  • Menthol cigarette smoking probably is a result of targeted marketing, not preference
Falling prevalence of US cigarette smoking
TUS-CPS NLMS mortality linkage presents unique opportunities for studies of mortality

- Detailed questions about range of tobacco products sold in the US
detailed usage patterns, lifetime history, cessation
- Nationally representative, so reflects US population and products used
wide age range, birth cohorts, and race/ethnicity group
US prevalence of different tobacco products

<table>
<thead>
<tr>
<th></th>
<th>Any</th>
<th>Cigs</th>
<th>Cigars</th>
<th>Hookah/pipe</th>
<th>E-cig</th>
<th>Smoke-less</th>
<th>≥2 types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>20.1%</td>
<td>15.1%</td>
<td>3.4%</td>
<td>1.2%</td>
<td>3.5%</td>
<td>2.3%</td>
<td>3.9%</td>
</tr>
<tr>
<td>High school</td>
<td>20.2%</td>
<td>8.0%</td>
<td>7.7%</td>
<td>6.2%</td>
<td>11.3%</td>
<td>5.8%</td>
<td>9.6%</td>
</tr>
</tbody>
</table>

- **Non-daily smokers**

NHIS, *MMWR*, 2017
Linked Mortality Data via National Longitudinal Mortality Study (NLMS)

- NLMS
  - Socioeconomic data collected as a part of CPS by US Census Bureau
  - Cause of death information from National Death Index through 2011
- Linked to a subset of TUS-CPS
- Linked mortality data currently available via RDC or Census
Mortality Analysis Example 1

Cigarette, Cigar, and Pipe Use and All-Cause and Cause-Specific Mortality, 1985-2011
Understanding the Health Risks of Non-cigarette Tobacco Product Use

- Because trends are recent, most epidemiologic studies lack users and detailed assessment.
Methods (1)

- 357,420 NLMS participants
- Exclusive use of cigarettes, cigars, and pipes
  - Ever users: Smoked 100+ cigarettes in the lifetime
  - Former/current; daily/non-daily (some day)
- All-cause mortality (n = 51,150)
- Cause-specific mortality
  - tobacco-related cancer; circulatory disease; cardiovascular disease; cerebrovascular disease; respiratory disease; diabetes
Methods (2)

- Cox proportional hazards proportional regression
  - Age as the underlying time metric
  - Covariates: sex, race/ethnicity, education, CPS survey year
  - Reference: never uses of cigarettes, cigars, pipes, or smokeless tobacco
- NLMS survey weights set to the non-institutional US population
Association of Exclusive Cigarette, Cigar, and Pipe Use with Tobacco-related Cancer Mortality, 1985-2011

<table>
<thead>
<tr>
<th></th>
<th>Exclusive cigarette</th>
<th></th>
<th>Exclusive cigar</th>
<th></th>
<th>Exclusive pipe</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-daily Daily</td>
<td>Non-daily Daily</td>
<td>Non-daily Daily</td>
<td>Non-daily Daily</td>
<td>Non-daily Daily</td>
<td>Non-daily Daily</td>
</tr>
<tr>
<td>Total, n</td>
<td>9,414 57,251</td>
<td>608 531</td>
<td>78 1,099</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death, n</td>
<td>236 2,984</td>
<td>* *</td>
<td>0 22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hazard Ratio</td>
<td>2.31 (2.01-2.65)</td>
<td>1.08 (0.45-2.61)</td>
<td>1.80 (1.20-2.69)</td>
<td>-</td>
<td>1.75 (1.16-2.64)</td>
<td></td>
</tr>
</tbody>
</table>

Christensen CH, et al. JAMA Intern Med 2018;178(4)
Mortality Analysis Example 2

All-cause and Cause-specific Mortality among Non-daily Cigarette Smokers, 1992-2011
National Health Interview Survey (NHIS)

- Annual household survey
- Nationally representative
- Tobacco data collected in Supplements since 1965
- Linkage to mortality data from National Death Index (NDI)

Total N = 70,913

<table>
<thead>
<tr>
<th></th>
<th>Daily smokers</th>
<th>Non-daily smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30 days</td>
<td>15 days</td>
</tr>
<tr>
<td></td>
<td>600 cigarettes</td>
<td>50 cigarettes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Daily</th>
<th>Lifelong Non-daily</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hazard Ratio</td>
<td>1.00 (reference)</td>
<td>2.50 (2.35-2.66)</td>
<td>1.72 (1.36-2.18)</td>
</tr>
</tbody>
</table>

Non-daily smokers lose 5 years on average of lifetime vs. never smokers: The NHIS 1991, 1992, and 1995
Limitations of previous analyses

- Lack of statistical power
detailed information on usage patterns
  - Number of cigarettes smoked per month
  - Reducing from daily to non-daily smoking
- Cause-specific mortality
Methods

- ~505,000 NLMS participants

Cigarette use

- Ever users: Smoked 100+ cigarettes in the lifetime
- Former/current (past 30 days); daily/non-daily (some day)
- Number of cigarettes smoked per day/on days they smoked
- Number of days smoked in the past 30 days
- Age started smoking; age at cessation; years since reducing from daily to non-daily smoking
All-cause Mortality by Daily and Non-daily Tobacco Use, 1992-2011

Manuscript in preparation; Poster #43, Session 4
All-cause Mortality by Years since Reducing from Daily to Non-daily Smoking, 1992-2011

Manuscript in preparation; Poster #43, Session 4
All-cause Mortality by Number of Cigarettes Smoked per Month among Daily Smokers, 1992-2011
All-cause Mortality by Number of Cigarettes Smoked per Month among Non-daily and Daily Smokers, 1992-2011
Summary

- Detailed tobacco usage data available in the harmonized TUS-CPS data set and linked mortality data are wonderful resources for mortality analyses.

- A large sample size of the harmonized TUS-CPS data enables examination of mortality risks in relation to detailed tobacco usage pattern.

- Risk estimates from the harmonized TUS-CPS data are representative of the U.S. civilian non-institutionalized adult population.
Acknowledgements

NCI/Division of Cancer Epidemiology and Genetics
  Barry I. Graubard
  Patricia Hartge
  Neil E. Caporaso

NCI/Division of Cancer Control and Population Sciences
  Carolyn Reyes-Guzman
  Anne Hartman
  Michele Bloch

FDA/Center for Tobacco Products
  Carol H. Christensen
  Brian Rostron
  Benjamin Apelberg

U.S. Census Bureau
  Candace M. Cosgrove
  Norman J. Johnson

IMS
  Timothy S. McNeel
  Todd Gibson
How to Use the Harmonized Data

James ‘Todd‘ Gibson
Information Management Services, Inc.
February 20, 2019
How to Use the Harmonized Data File

- Obtaining the Data
- Harmonized Data File Examples
- Hands-on Exercises
- Questions
- Contact Information
Obtaining the Data
Obtaining the Data

- Data available on the TUS-CPS website: https://cancercontrol.cancer.gov/brp/tcrb/tus-cps
Obtaining the Data

- 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 in the dataset. Program is called by main program.

- Other Helpful Information
  - Technical Documentation: Overview of CPS, TUS and Harmonized Data
  - Proc Contents of the Data File
  - Unweighted Frequency Tables of All Variables
  - Microsoft Excel Table Listing the TUS Harmonized Variables and the Source Variable by Survey Wave.
Harmonized Data File Examples
Harmonized Data File Examples

- **Example 1:**
  - Example of a variable that was easy to harmonize.

- **Example 2:**
  - More complex harmonized variable.

- **Example 3:**
  - Example of a variable constructed from multiple variables.
Harmonized Data File Examples

- Example 1: Example of a variable that was easy to harmonize
  - Harmonized Variable CIG100: Smoked 100 cigarettes in entire life
  - Supplement question: Have you smoked at least 100 cigarettes in your entire life?
  - Variable names and column positions changed over time in individual survey files. User would need to read in the 26 survey files and create a new variable. 1992-93 variable coded differently than other years.
    - 1992-93: A_S32 (col 361)
    - 1995-96: PES32 (815-816)
    - 1998-99: PES32 (857-858)
    - 2000, 2001-02: PES 32 (859-860)
    - 2003: PEA1 (879-880)
Harmonized Data File Examples

- Example 1:
  - SAS program code to generate unweighted table of CIG100 by survey wave

```sas
%Data Harmon;
    Set MyData.Harmon;
    Keep SurWave Cig100 SRWeight;
Run;

%Proc Tabulate Data=Harmon Missing;
    Title1 "Tobacco Use Supplement to The Current Population Survey";
    Title2 "Harmonized Data File";
    Title3 "Variable: CIG100";
    Class SurWave Cig100;
    Table Cig100 All="Total",
        SurWave=""*N="Sample"*F=Comma8;.
Run;
```
# Harmonized Data File Examples

## Tobacco Use Supplement to The Current Population Survey

### Harmonized Data File

#### Variable: CIG100

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoked 100 cigarettes in entire life</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-9: No response</td>
<td>43</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>2</td>
<td>1</td>
<td>.</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>-3: Refused</td>
<td>.</td>
<td>59</td>
<td>76</td>
<td>53</td>
<td>82</td>
<td>236</td>
<td>174</td>
<td>174</td>
<td>171</td>
</tr>
<tr>
<td>-2: Don’t know</td>
<td>625</td>
<td>160</td>
<td>241</td>
<td>179</td>
<td>242</td>
<td>212</td>
<td>237</td>
<td>222</td>
<td>183</td>
</tr>
<tr>
<td>1: Yes</td>
<td>110,998</td>
<td>88,179</td>
<td>80,557</td>
<td>56,244</td>
<td>81,935</td>
<td>74,701</td>
<td>70,315</td>
<td>63,455</td>
<td>58,263</td>
</tr>
<tr>
<td>Total</td>
<td>228,884</td>
<td>186,922</td>
<td>176,135</td>
<td>124,350</td>
<td>185,242</td>
<td>183,361</td>
<td>171,602</td>
<td>170,968</td>
<td>163,562</td>
</tr>
</tbody>
</table>

---

**NATIONAL CANCER INSTITUTE**
Harmonized Data File Examples

- Example 2: More complex harmonized variable.
  - Harmonized Variable CIGTYPE: Type of cigarette
    - 9: No response
    - 3: Refused
    - 2: Don’t know
    - 1: Menthol
    - 2: Non-menthol
    - 3: No usual type
    - 1: Not in universe
  - Every day and some day smokers, 2003 and later
  - Supplement questions:
    - 2003, 2006-07: Is your usual cigarette brand menthol or non-menthol?
    - 2010-11, 2014-15: Do you usually smoke menthol or non-menthol cigarettes?
  - Two separate questions
    - B2 (Every day smokers), C2 (Some day smokers)
Harmonized Data File Examples

- Example 2:
  - SAS program code to generate unweighted table of CIGTYPE by survey wave

```sas
Data Harmon;
  Set MyData.Harmon;
  Keep SurWave CigType SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to The Current Population Survey";
  Title2 "Harmonized Data File";
  Title3 "Variable: CigType";
  Class SurWave CigType;
  Table CigType All="Total",
      SurWave="*N="Sample"*F=Comma8.;
Run;
```
### Harmonized Data File Examples

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

**Harmonized Data File**

**Variable: CigType**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>.</td>
<td>228,552</td>
<td>187,141</td>
<td>176,452</td>
<td>124,582</td>
<td>185,568</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.9: No response</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>.3: Refused</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>323</td>
<td>190</td>
<td>257</td>
</tr>
<tr>
<td>.2: Don’t know</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>107</td>
<td>94</td>
<td>86</td>
</tr>
<tr>
<td>.1: Not in universe</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>149,166</td>
<td>140,522</td>
<td>143,754</td>
</tr>
<tr>
<td>1: Menthol</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>8,576</td>
<td>7,718</td>
<td>7,624</td>
</tr>
<tr>
<td>2: Non-menthol</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>24,441</td>
<td>22,458</td>
<td>18,758</td>
</tr>
<tr>
<td>3: No usual type</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>1,087</td>
<td>941</td>
<td>807</td>
</tr>
<tr>
<td>Total</td>
<td>228,552</td>
<td>187,141</td>
<td>176,452</td>
<td>124,582</td>
<td>185,568</td>
<td>183,810</td>
<td>172,023</td>
<td>171,366</td>
<td>163,920</td>
</tr>
</tbody>
</table>
Example 3: Variable constructed from multiple variables.

- Harmonized Variable WORKBAN: Smoking policy at work
  - -9: indeterminate
  - -1: not in universe
  - 1: Not allowed in any indoor public or work areas
  - 2: Allowed in some/all indoor public or work areas or not applicable
  - 3: No policy

- Not available in January and May 2000

- Constructed from:
  - WKPOLICY: Does your place of work have an official policy that restricts smoking in any way?
  - WKPOLIND: Smoking policy for indoor public or common areas such as lobbies, rest rooms and lunch rooms.
  - WKPOLWRK: Smoking policy for work areas
Harmonized Data File Examples

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-9: No response</td>
<td>2,469</td>
<td>548</td>
<td>712</td>
<td>981</td>
<td>20</td>
<td>61</td>
<td>50</td>
<td>56</td>
</tr>
<tr>
<td>-3: Refused</td>
<td>.</td>
<td>197</td>
<td>255</td>
<td>338</td>
<td>80</td>
<td>105</td>
<td>106</td>
<td>89</td>
</tr>
<tr>
<td>-2: Don't know</td>
<td>1,757</td>
<td>534</td>
<td>594</td>
<td>760</td>
<td>542</td>
<td>662</td>
<td>265</td>
<td>285</td>
</tr>
<tr>
<td>-1: Not in universe</td>
<td>124,958</td>
<td>100,825</td>
<td>93,194</td>
<td>98,072</td>
<td>105,190</td>
<td>98,378</td>
<td>96,448</td>
<td>94,071</td>
</tr>
<tr>
<td>1: Yes</td>
<td>81,178</td>
<td>73,223</td>
<td>71,378</td>
<td>74,323</td>
<td>70,993</td>
<td>65,279</td>
<td>70,085</td>
<td>65,016</td>
</tr>
<tr>
<td>2: No</td>
<td>18,190</td>
<td>11,814</td>
<td>10,319</td>
<td>11,094</td>
<td>6,985</td>
<td>7,538</td>
<td>4,411</td>
<td>4,403</td>
</tr>
<tr>
<td>Total</td>
<td>228,652</td>
<td>187,141</td>
<td>176,452</td>
<td>186,568</td>
<td>183,810</td>
<td>172,023</td>
<td>171,365</td>
<td>163,920</td>
</tr>
</tbody>
</table>
Harmonized Data File Examples

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9: No response</td>
<td>405</td>
<td>5</td>
<td>9</td>
<td>8</td>
<td>2</td>
<td>9</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>3: Refused</td>
<td>.</td>
<td>17</td>
<td>10</td>
<td>14</td>
<td>19</td>
<td>21</td>
<td>46</td>
<td>42</td>
</tr>
<tr>
<td>2: Don't know (not read)</td>
<td>412</td>
<td>139</td>
<td>103</td>
<td>137</td>
<td>105</td>
<td>113</td>
<td>138</td>
<td>152</td>
</tr>
<tr>
<td>1: Not in universe</td>
<td>147,374</td>
<td>113,918</td>
<td>105,074</td>
<td>111,245</td>
<td>112,817</td>
<td>106,744</td>
<td>101,280</td>
<td>98,904</td>
</tr>
<tr>
<td>1: Not allowed in any public areas</td>
<td>49,661</td>
<td>56,847</td>
<td>58,856</td>
<td>62,679</td>
<td>62,020</td>
<td>56,828</td>
<td>62,468</td>
<td>56,579</td>
</tr>
<tr>
<td>2: Allowed in some public areas</td>
<td>28,035</td>
<td>14,288</td>
<td>10,654</td>
<td>10,091</td>
<td>7,916</td>
<td>7,147</td>
<td>6,336</td>
<td>6,540</td>
</tr>
<tr>
<td>3: Allowed in all public areas</td>
<td>2,064</td>
<td>1,668</td>
<td>1,378</td>
<td>1,203</td>
<td>765</td>
<td>1,004</td>
<td>687</td>
<td>1,214</td>
</tr>
<tr>
<td>4: Not applicable (not read)</td>
<td>601</td>
<td>259</td>
<td>152</td>
<td>191</td>
<td>166</td>
<td>157</td>
<td>393</td>
<td>476</td>
</tr>
<tr>
<td>Total</td>
<td>228,552</td>
<td>187,141</td>
<td>176,452</td>
<td>186,568</td>
<td>183,810</td>
<td>172,023</td>
<td>171,365</td>
<td>163,920</td>
</tr>
</tbody>
</table>
## Harmonized Data File Examples

### Tobacco Use Supplement to The Current Population Survey

#### Harmonized Data File

**Variable: WKPOLWRK**

**Unweighted Counts**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-9: No response</td>
<td>479</td>
<td>9</td>
<td>13</td>
<td>11</td>
<td>6</td>
<td>14</td>
<td>20</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>-3: Refused</td>
<td>..</td>
<td>21</td>
<td>14</td>
<td>18</td>
<td>19</td>
<td>26</td>
<td>50</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>-2: Don’t know (not read)</td>
<td>232</td>
<td>76</td>
<td>83</td>
<td>102</td>
<td>90</td>
<td>75</td>
<td>81</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>-1: Not in universe</td>
<td>147,374</td>
<td>113,918</td>
<td>105,074</td>
<td>111,245</td>
<td>112,817</td>
<td>106,744</td>
<td>101,280</td>
<td>98,904</td>
<td></td>
</tr>
<tr>
<td>1: Not allowed in any work areas</td>
<td>66,034</td>
<td>64,351</td>
<td>64,525</td>
<td>67,910</td>
<td>66,870</td>
<td>61,192</td>
<td>67,477</td>
<td>62,045</td>
<td></td>
</tr>
<tr>
<td>2: Allowed in some work areas</td>
<td>13,193</td>
<td>7,314</td>
<td>5,469</td>
<td>5,120</td>
<td>3,436</td>
<td>3,061</td>
<td>1,841</td>
<td>1,815</td>
<td></td>
</tr>
<tr>
<td>3: Allowed in all work areas</td>
<td>1,091</td>
<td>1,336</td>
<td>1,172</td>
<td>1,021</td>
<td>482</td>
<td>815</td>
<td>349</td>
<td>723</td>
<td></td>
</tr>
<tr>
<td>4: Not applicable (not read)</td>
<td>149</td>
<td>116</td>
<td>102</td>
<td>133</td>
<td>90</td>
<td>96</td>
<td>267</td>
<td>292</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>228,552</td>
<td>187,141</td>
<td>176,452</td>
<td>185,568</td>
<td>183,810</td>
<td>172,023</td>
<td>171,365</td>
<td>163,920</td>
<td></td>
</tr>
<tr>
<td>WORKBAN</td>
<td>WKPOLICY</td>
<td>WKPOLIND</td>
<td>WKPOLWRK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3: Refused</td>
<td>-3: Refused</td>
<td>-3: Refused</td>
<td>-3: Refused</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2: Don’t know</td>
<td>-2: Don’t know</td>
<td>-2: Don’t know</td>
<td>-2: Don’t know</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Not allowed in any public or work areas</td>
<td>1: Yes</td>
<td>1: Not allowed in any public areas</td>
<td>1: Not allowed in any work areas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2: Allowed in some/all indoor public or work areas or not applicable</td>
<td>1: Yes</td>
<td>2: Allowed in some public areas</td>
<td>2: Allowed in some work areas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: No policy</td>
<td>2: No</td>
<td>-1: Not in universe</td>
<td>-1: Not in universe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

WKPOLICY: Does your place of work have an official policy that restricts smoking in any way
WKPOLIND: Smoking policy for indoor public or common areas such as lobbies, rest rooms and lunch rooms.
WKPOLWRK: Smoking policy for work areas
# Harmonized Data File Examples

## Tobacco Use Supplement to The Current Population Survey

**Harmonized Data File**

**Variable: WORKBAN**

**Unweighted Counts**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>-9: Indeterminate</td>
<td>5,279</td>
<td>1,480</td>
<td>1,737</td>
<td>2,286</td>
<td>816</td>
<td>1,005</td>
<td>665</td>
<td>681</td>
</tr>
<tr>
<td>-1: Not in universe</td>
<td>124,968</td>
<td>100,825</td>
<td>93,194</td>
<td>98,072</td>
<td>105,190</td>
<td>98,378</td>
<td>96,443</td>
<td>94,071</td>
</tr>
<tr>
<td>1: Not allowed in any indoor public or work areas</td>
<td>45,898</td>
<td>64,560</td>
<td>57,006</td>
<td>61,046</td>
<td>60,799</td>
<td>55,720</td>
<td>61,724</td>
<td>55,884</td>
</tr>
<tr>
<td>2: Allowed in some/all indoor public or work areas or not applicable</td>
<td>34,227</td>
<td>18,462</td>
<td>14,156</td>
<td>13,071</td>
<td>10,020</td>
<td>9,332</td>
<td>8,117</td>
<td>8,881</td>
</tr>
<tr>
<td>3: No policy</td>
<td>18,180</td>
<td>11,814</td>
<td>10,319</td>
<td>11,094</td>
<td>6,986</td>
<td>7,538</td>
<td>4,411</td>
<td>4,403</td>
</tr>
<tr>
<td>Total</td>
<td>228,552</td>
<td>187,141</td>
<td>176,452</td>
<td>185,568</td>
<td>183,810</td>
<td>172,023</td>
<td>171,365</td>
<td>163,920</td>
</tr>
</tbody>
</table>
Hands-on Exercises
Hands-on Exercises

- Exercise 1: Creating a SAS dataset from the ASCII text file
- Exercise 2: Weights
  - Current cigarette smoking status by survey: CIGSTAT * SurWave
- Exercise 3:
  - Home smoking rules by survey and geographic region: Region * SurWave * HMSMKPOL
- Exercise 4:
  - Ever cigar use by survey and sex: SurWave * Sex * EVERCIGR
Hands-on Exercises

Exercise 1: Creating a SAS dataset from the ASCII text file

- Download the data and SAS code from TUS-CPS website.
- Unzip the data file
- Modify `<directory>` in the main SAS program to reflect where the format program and data are.
  - Libname MyLib "<directory>";
  - %Include "<directory>harmonzd.tus_cps.1992.through.2015.vers_0_9.beta.formats.sas";
- Run SAS program to permanent SAS dataset.
Hands-on Exercises

- Exercise 2: Working with Weights
  - Self-response weight needs divided by the number of surveys in the survey wave.
    - There were 2 surveys in 2000.
    - All other survey waves had 3.
  - Generate tables of current cigarette status by survey wave
    - Table of weighted counts
    - Table of weighted percentages
Hands-on Exercises

LibName MyData "<directory>";
Include "<directory>\harmoniz.tus_ops.1992.through.2015.vers_0_9.beta.formats.cas";

Data Harmon;
  Set MyData.Harmon;
  If SuzWave=4 Then SRWeight=SRWeight/2;
  Else SRWeight=SRWeight/3;
  Keep SuzWave CIGSTAT SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to The Current Population Survey";
  Title2 "Harmonized Data File";
  Title3 "Variable: CIGSTAT";
  Title4 "Weighted Counts";
  Class SuzWave CIGSTAT;
  Var SRWeight;
  Table CIGSTAT All="Total",
    SuzWave=""*SRWeight=""*Sum="Population"*F=Comma12.:
Run;

Proc Tabulate Data=Harmon Missing;
  Title4 "Weighted Percentages";
  Class SuzWave CIGSTAT;
  Var SRWeight;
  Table CIGSTAT All="Total",
    SuzWave=""*SRWeight=""*PctSum<CIGSTAT All>="Percent"*F=9.2:
Run;
### Hands-on Exercises

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

Harmonized Data File

**Variable: CIGSTAT**

**Weighted Counts**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>691,028</td>
<td>468,256</td>
<td>662,677</td>
<td>615,466</td>
<td>732,705</td>
<td>912,321</td>
<td>977,951</td>
<td>1,185,689</td>
<td>1,255,569</td>
</tr>
<tr>
<td>-9: Indeterminate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1: Never</td>
<td>96,916,000</td>
<td>102,859,579</td>
<td>109,515,348</td>
<td>111,861,687</td>
<td>117,157,403</td>
<td>129,048,509</td>
<td>135,038,811</td>
<td>149,368,484</td>
<td>162,113,749</td>
</tr>
<tr>
<td>2: Every day</td>
<td>37,301,984</td>
<td>37,046,026</td>
<td>35,369,907</td>
<td>34,713,580</td>
<td>34,142,310</td>
<td>32,227,027</td>
<td>32,494,777</td>
<td>28,933,246</td>
<td>25,474,113</td>
</tr>
<tr>
<td>3: Some days</td>
<td>8,241,594</td>
<td>8,066,711</td>
<td>8,280,345</td>
<td>9,116,017</td>
<td>8,568,705</td>
<td>7,740,653</td>
<td>8,111,173</td>
<td>7,817,083</td>
<td>7,453,035</td>
</tr>
<tr>
<td>Total</td>
<td>186,049,931</td>
<td>192,057,570</td>
<td>196,489,114</td>
<td>201,105,905</td>
<td>204,549,853</td>
<td>212,415,094</td>
<td>220,631,534</td>
<td>229,791,812</td>
<td>241,120,556</td>
</tr>
</tbody>
</table>
### Hands-on Exercises

| Tobacco Use Supplement to The Current Population Survey |
| Harmonized Data File |
| Variable: CIGSTAT |
| Weighted Percentages |

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
</tr>
<tr>
<td><strong>-9: Indeterminate</strong></td>
<td>0.37</td>
<td>0.24</td>
<td>0.33</td>
<td>0.31</td>
<td>0.36</td>
<td>0.43</td>
<td>0.44</td>
<td>0.52</td>
<td>0.52</td>
</tr>
<tr>
<td><strong>1: Never</strong></td>
<td>51.87</td>
<td>53.56</td>
<td>55.17</td>
<td>55.62</td>
<td>57.28</td>
<td>60.75</td>
<td>61.21</td>
<td>65.00</td>
<td>67.23</td>
</tr>
<tr>
<td><strong>2: Every day</strong></td>
<td>19.96</td>
<td>19.29</td>
<td>17.81</td>
<td>17.26</td>
<td>16.69</td>
<td>15.17</td>
<td>14.73</td>
<td>12.59</td>
<td>10.56</td>
</tr>
<tr>
<td><strong>3: Some days</strong></td>
<td>4.41</td>
<td>4.20</td>
<td>4.17</td>
<td>4.53</td>
<td>4.19</td>
<td>3.64</td>
<td>3.68</td>
<td>3.40</td>
<td>3.09</td>
</tr>
<tr>
<td><strong>4: Former</strong></td>
<td>23.39</td>
<td>22.71</td>
<td>22.51</td>
<td>22.28</td>
<td>21.49</td>
<td>20.00</td>
<td>19.95</td>
<td>18.49</td>
<td>18.59</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Hands-on Exercises

square Exercise 3:

• Home smoking rules by survey and geographic region: Region * SurWave * HSMKPOL

• Not available in January and May 2000

• Generate tables showing weighed percentages, sample counts and population.
Hands-on Exercises

LibName MyData "<directory>";
%Include "<directory>\harmonzd.tus_cps.1992.through.2015.vers_0_9.betaformats.sas";

Data Harmon;
  Set MyData.Harmon;
  If SurWave=4 Then SRWeight=SRWeight/2;
  Else SRWeight=SRWeight/3;
  Keep SurWave Region HMSMKPOL SRWeight;
Run;

Proc Tabulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to The Current Population Survey";
  Title2 "Harmonized Data File";
  Title3 "Variable: HMSMKPOL";
  Where SurWave In(1,2,3);
  Class SurWave Region HMSMKPOL;
  Var SRWeight;
  Table SurWave=""*(All="Total" Region="" ),
           HMSMKPOL*(SRWeight=""*PctSum<HMSMKPOL>="Percent"*F=7.2)
           All="Total"*(N="Sample"*F=Comma8. SRWeight=""*Sum="Population"*F=Comma12.);
Run;
# Hands-on Exercises

## Tobacco Use Supplement to The Current Population Survey

*Harmonized Data File*

*Variable: HMSMKPOL*

<table>
<thead>
<tr>
<th>Home smoking rules</th>
<th>9: No response</th>
<th>3: Refused</th>
<th>2: Don’t know</th>
<th>1: No one is allowed to smoke anywhere</th>
<th>2: Smoking is allowed in some places or at some times</th>
<th>3: Smoking is allowed anywhere</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Sample</td>
</tr>
<tr>
<td>1992-1993</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.70</td>
<td>.</td>
<td>.</td>
<td>42.79</td>
<td>25.82</td>
<td>30.69</td>
<td>228,562</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>0.77</td>
<td>.</td>
<td>.</td>
<td>40.61</td>
<td>29.31</td>
<td>29.31</td>
<td>53,254</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.70</td>
<td>.</td>
<td>.</td>
<td>35.96</td>
<td>28.57</td>
<td>34.77</td>
<td>56,533</td>
</tr>
<tr>
<td>3: South</td>
<td>0.84</td>
<td>.</td>
<td>.</td>
<td>40.69</td>
<td>24.31</td>
<td>34.35</td>
<td>68,273</td>
</tr>
<tr>
<td>4: West</td>
<td>0.72</td>
<td>.</td>
<td>.</td>
<td>55.78</td>
<td>21.89</td>
<td>21.60</td>
<td>48,492</td>
</tr>
<tr>
<td>Total</td>
<td>0.31</td>
<td>0.30</td>
<td>0.29</td>
<td>52.82</td>
<td>22.03</td>
<td>24.24</td>
<td>187,141</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>0.44</td>
<td>0.42</td>
<td>0.41</td>
<td>49.91</td>
<td>25.50</td>
<td>23.33</td>
<td>41,039</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.24</td>
<td>0.26</td>
<td>0.26</td>
<td>45.75</td>
<td>24.88</td>
<td>26.61</td>
<td>46,527</td>
</tr>
<tr>
<td>3: South</td>
<td>0.25</td>
<td>0.30</td>
<td>0.24</td>
<td>51.48</td>
<td>20.79</td>
<td>26.94</td>
<td>57,167</td>
</tr>
<tr>
<td>4: West</td>
<td>0.36</td>
<td>0.26</td>
<td>0.30</td>
<td>65.27</td>
<td>17.80</td>
<td>16.02</td>
<td>42,468</td>
</tr>
<tr>
<td>Total</td>
<td>0.44</td>
<td>0.42</td>
<td>0.41</td>
<td>60.31</td>
<td>19.24</td>
<td>19.19</td>
<td>176,462</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>0.65</td>
<td>0.72</td>
<td>0.60</td>
<td>59.24</td>
<td>21.20</td>
<td>18.59</td>
<td>36,322</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.36</td>
<td>0.31</td>
<td>0.25</td>
<td>53.85</td>
<td>21.84</td>
<td>23.39</td>
<td>42,983</td>
</tr>
<tr>
<td>3: South</td>
<td>0.42</td>
<td>0.32</td>
<td>0.36</td>
<td>59.01</td>
<td>18.69</td>
<td>21.18</td>
<td>54,314</td>
</tr>
<tr>
<td>4: West</td>
<td>0.37</td>
<td>0.41</td>
<td>0.43</td>
<td>70.95</td>
<td>15.69</td>
<td>12.16</td>
<td>42,833</td>
</tr>
</tbody>
</table>
# Hands-on Exercises

## Tobacco Use Supplement to The Current Population Survey

### Harmonized Data File

**Variable:** HMSMKPOL

<table>
<thead>
<tr>
<th>Home smoking rules</th>
<th>3: No response</th>
<th>2: Refused</th>
<th>1: Don’t know</th>
<th>1: No one is allowed to smoke anywhere</th>
<th>2: Smoking is allowed in some places or at some times</th>
<th>3: Smoking is allowed anywhere</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Sample</td>
</tr>
<tr>
<td><strong>2001-2002</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>165,668</td>
</tr>
<tr>
<td>Total</td>
<td>0.38</td>
<td>0.51</td>
<td>0.51</td>
<td>66.28</td>
<td>17.11</td>
<td>15.21</td>
<td>165,668</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>0.47</td>
<td>0.91</td>
<td>0.79</td>
<td>64.52</td>
<td>18.61</td>
<td>14.49</td>
<td>39,906</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.32</td>
<td>0.45</td>
<td>0.47</td>
<td>68.30</td>
<td>19.48</td>
<td>18.97</td>
<td>48,140</td>
</tr>
<tr>
<td>3: South</td>
<td>0.36</td>
<td>0.41</td>
<td>0.43</td>
<td>65.07</td>
<td>16.52</td>
<td>17.20</td>
<td>64,026</td>
</tr>
<tr>
<td>4: West</td>
<td>0.37</td>
<td>0.40</td>
<td>0.44</td>
<td>75.72</td>
<td>14.19</td>
<td>8.68</td>
<td>44,406</td>
</tr>
<tr>
<td><strong>2003</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>183,810</td>
</tr>
<tr>
<td>Total</td>
<td>0.67</td>
<td>0.72</td>
<td>0.42</td>
<td>72.80</td>
<td>12.31</td>
<td>13.08</td>
<td>38,623</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>0.85</td>
<td>0.97</td>
<td>0.54</td>
<td>78.94</td>
<td>14.10</td>
<td>12.69</td>
<td>47,266</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.49</td>
<td>0.80</td>
<td>0.50</td>
<td>65.69</td>
<td>15.68</td>
<td>16.64</td>
<td>75,495,450</td>
</tr>
<tr>
<td>3: South</td>
<td>0.69</td>
<td>0.56</td>
<td>0.36</td>
<td>72.29</td>
<td>11.92</td>
<td>14.17</td>
<td>53,081</td>
</tr>
<tr>
<td>4: West</td>
<td>0.68</td>
<td>0.66</td>
<td>0.34</td>
<td>62.36</td>
<td>7.98</td>
<td>7.96</td>
<td>44,406</td>
</tr>
<tr>
<td><strong>2006-2007</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>172,023</td>
</tr>
<tr>
<td>Total</td>
<td>0.66</td>
<td>0.66</td>
<td>0.37</td>
<td>77.80</td>
<td>10.35</td>
<td>10.17</td>
<td>42,016</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>0.91</td>
<td>0.74</td>
<td>0.36</td>
<td>76.08</td>
<td>12.47</td>
<td>9.44</td>
<td>33,491</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.54</td>
<td>0.69</td>
<td>0.37</td>
<td>72.33</td>
<td>12.80</td>
<td>13.26</td>
<td>49,150,326</td>
</tr>
<tr>
<td>3: South</td>
<td>0.64</td>
<td>0.61</td>
<td>0.36</td>
<td>76.87</td>
<td>10.18</td>
<td>11.34</td>
<td>54,428</td>
</tr>
<tr>
<td>4: West</td>
<td>0.60</td>
<td>0.64</td>
<td>0.38</td>
<td>86.00</td>
<td>6.49</td>
<td>5.69</td>
<td>50,398,371</td>
</tr>
</tbody>
</table>
# Hands-on Exercises

## Tobacco Use Supplement to The Current Population Survey

### Harmonized Data File

**Variable:** HSMSMKPOL

### Home smoking rules

<table>
<thead>
<tr>
<th></th>
<th>-9: No response</th>
<th>-3: Refused</th>
<th>-2: Don't know</th>
<th>1: No one is allowed to smoke anywhere</th>
<th>2: Smoking is allowed in some places or at some times</th>
<th>3: Smoking is allowed anywhere</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Percent</td>
<td>Sample</td>
</tr>
<tr>
<td><strong>2010-2011</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.96</td>
<td>0.91</td>
<td>0.28</td>
<td>82.13</td>
<td>8.91</td>
<td>7.71</td>
<td>171,365</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>0.92</td>
<td>0.91</td>
<td>0.24</td>
<td>81.18</td>
<td>9.18</td>
<td>7.58</td>
<td>34,631</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.86</td>
<td>1.24</td>
<td>0.36</td>
<td>77.18</td>
<td>10.16</td>
<td>10.20</td>
<td>41,654</td>
</tr>
<tr>
<td>3: South</td>
<td>0.98</td>
<td>0.80</td>
<td>0.24</td>
<td>81.21</td>
<td>8.20</td>
<td>8.57</td>
<td>54,251</td>
</tr>
<tr>
<td>4: West</td>
<td>1.05</td>
<td>0.78</td>
<td>0.29</td>
<td>89.01</td>
<td>4.75</td>
<td>4.12</td>
<td>40,829</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0.89</td>
<td>0.97</td>
<td>0.21</td>
<td>84.70</td>
<td>6.18</td>
<td>7.05</td>
<td>163,920</td>
</tr>
<tr>
<td>1: Northeast</td>
<td>1.02</td>
<td>1.10</td>
<td>0.23</td>
<td>83.86</td>
<td>6.95</td>
<td>6.84</td>
<td>28,370</td>
</tr>
<tr>
<td>2: Midwest</td>
<td>0.76</td>
<td>1.20</td>
<td>0.26</td>
<td>81.68</td>
<td>7.81</td>
<td>8.50</td>
<td>35,894</td>
</tr>
<tr>
<td>3: South</td>
<td>0.81</td>
<td>0.81</td>
<td>0.17</td>
<td>82.84</td>
<td>6.53</td>
<td>7.83</td>
<td>58,602</td>
</tr>
<tr>
<td>4: West</td>
<td>1.02</td>
<td>0.89</td>
<td>0.22</td>
<td>89.44</td>
<td>3.75</td>
<td>4.66</td>
<td>41,054</td>
</tr>
</tbody>
</table>
Hands-on Exercises

Exercise 4:

• Ever used cigars by sex and survey wave: SurWave * Sex * CIGREVER

• Create a SurWave recode variable that combines September 1998 data with January and May 2000

• Generate tables showing weighed percentages, sample counts and population.
Hands-on Exercises

LibName MyData "<directory>";
%Include "<directory>\harmonzd.tus_cps.1992.through.2015.vers_0.9.betasformats.sas";

Proc Format;
  Value SurWaveRF
    1 = "1992-1993"
    2 = "1995-1996"
    3 = "1998"
    4 = "1998,2000"
    5 = "2001-2002"
    6 = "2003"
    7 = "2004-2007"
    8 = "2010-2011"
    9 = "2014-2015"
  ;

Data Harmon;
  Set MyData.Harmon;
  If SurYear=1998 Then SurWaveR=4;
  Else SurWaveR=SurWave;
  If SurWave=3 Then SRSWeight=SRSWeight/2;
  Else SRSWeight=SRSWeight/3;
  Attrib SurWaveR Label = "Survey wave recode" Format=SurWaveRF.;
  Keep SurYear SurMonth SurWave SurWaveR SEX CIGREVER PIPEVER SRSWeight;
Run;

Proc Tabbulate Data=Harmon Missing;
  Title1 "Tobacco Use Supplement to The Current Population Survey";
  Title2 "Harmonized Data File";
  Title3 "Variable: CIGREVER (Ever used cigars except cigarette or little filtered ciger)"
  Title4 "No Response, Refused and Don’t Know Excluded"
  Where CIGREVER Not in (-9,-9,-9);
  Class SurWave Sex CIGREVER;
  Var SRSWeight;
  Table SurWaveR=""
    ([All="Total", Sex=" ")(CIGREVER=""|SRSWeight=""|FctSum<CIGREVER>=""Percent=""|F=7.2)
    All="Total" *(N="Sample"|F=Comma8. SRSWeight=""|Sum=Population""|F=Comma12.));
Run;
# Hands-on Exercises

## Tobacco Use Supplement to The Current Population Survey

**Harmonized Data File**

Variable: CIGREVER (Ever used cigars <regular cigar, cigarillos or little filtered cigars>?)

No Response, Refused and Don’t Know Excluded

<table>
<thead>
<tr>
<th>Year</th>
<th>Total 1: Yes Percent</th>
<th>Total 2: No Percent</th>
<th>Total Sample</th>
<th>Total Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992-1993</td>
<td>3.96</td>
<td>96.04</td>
<td>227,158</td>
<td>185,726,105</td>
</tr>
<tr>
<td>1995-1996</td>
<td>3.64</td>
<td>96.36</td>
<td>186,448</td>
<td>191,310,283</td>
</tr>
<tr>
<td>1999</td>
<td>3.75</td>
<td>96.25</td>
<td>113,695</td>
<td>197,950,668</td>
</tr>
<tr>
<td>1998-2000</td>
<td>10.62</td>
<td>89.38</td>
<td>185,970</td>
<td>231,956,057</td>
</tr>
<tr>
<td>2001-2002</td>
<td>13.40</td>
<td>86.60</td>
<td>184,562</td>
<td>203,403,835</td>
</tr>
<tr>
<td>2003</td>
<td>11.72</td>
<td>88.28</td>
<td>181,659</td>
<td>209,848,309</td>
</tr>
<tr>
<td>2006-2007</td>
<td>22.64</td>
<td>77.36</td>
<td>170,264</td>
<td>218,264,724</td>
</tr>
<tr>
<td>2010-2011</td>
<td>16.48</td>
<td>83.52</td>
<td>168,764</td>
<td>226,527,076</td>
</tr>
<tr>
<td>2014-2015</td>
<td>17.27</td>
<td>82.73</td>
<td>151,867</td>
<td>238,103,106</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Total 1: Yes Percent</th>
<th>Total 2: No Percent</th>
<th>Total Sample</th>
<th>Total Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992-1993</td>
<td>7.97</td>
<td>92.03</td>
<td>97,195</td>
<td>88,543,365</td>
</tr>
<tr>
<td>1995-1996</td>
<td>7.31</td>
<td>92.69</td>
<td>76,650</td>
<td>91,284,565</td>
</tr>
<tr>
<td>1999</td>
<td>7.37</td>
<td>92.63</td>
<td>49,037</td>
<td>94,550,565</td>
</tr>
<tr>
<td>1998-2000</td>
<td>19.78</td>
<td>80.22</td>
<td>80,552</td>
<td>111,006,247</td>
</tr>
<tr>
<td>2001-2002</td>
<td>23.50</td>
<td>76.50</td>
<td>80,514</td>
<td>97,345,577</td>
</tr>
<tr>
<td>2003</td>
<td>20.46</td>
<td>79.54</td>
<td>78,813</td>
<td>100,429,209</td>
</tr>
<tr>
<td>2006-2007</td>
<td>37.20</td>
<td>62.80</td>
<td>73,832</td>
<td>105,032,779</td>
</tr>
<tr>
<td>2010-2011</td>
<td>27.03</td>
<td>72.97</td>
<td>74,341</td>
<td>109,021,609</td>
</tr>
<tr>
<td>2014-2015</td>
<td>28.21</td>
<td>71.79</td>
<td>72,256</td>
<td>114,248,135</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Total 1: Yes Percent</th>
<th>Total 2: No Percent</th>
<th>Total Sample</th>
<th>Total Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992-1993</td>
<td>0.30</td>
<td>99.70</td>
<td>129,963</td>
<td>97,082,750</td>
</tr>
<tr>
<td>1995-1996</td>
<td>0.28</td>
<td>99.72</td>
<td>107,798</td>
<td>100,025,778</td>
</tr>
<tr>
<td>1999</td>
<td>0.45</td>
<td>99.55</td>
<td>64,858</td>
<td>103,408,103</td>
</tr>
<tr>
<td>1998-2000</td>
<td>2.20</td>
<td>97.80</td>
<td>105,318</td>
<td>120,948,810</td>
</tr>
<tr>
<td>2001-2002</td>
<td>4.14</td>
<td>95.86</td>
<td>103,948</td>
<td>105,056,258</td>
</tr>
<tr>
<td>2003</td>
<td>3.71</td>
<td>96.29</td>
<td>102,846</td>
<td>109,419,100</td>
</tr>
<tr>
<td>2006-2007</td>
<td>9.13</td>
<td>90.87</td>
<td>96,432</td>
<td>113,231,944</td>
</tr>
<tr>
<td>2010-2011</td>
<td>6.68</td>
<td>93.32</td>
<td>94,423</td>
<td>117,505,267</td>
</tr>
<tr>
<td>2014-2015</td>
<td>7.17</td>
<td>92.83</td>
<td>89,601</td>
<td>123,854,971</td>
</tr>
</tbody>
</table>
Contacts

- For general TUS questions: ncidccpsbrpadvances@mail.nih.gov
- For TUS harmonized data questions: Email Carolyn Reyes-Guzman (carolyn.reyes-guzman@nih.gov)
- For weighting or TUS variable-specific questions: Email Todd Gibson (gibsont@imsweb.com) and Dr. Reyes-Guzman (carolyn.reyes-guzman@nih.gov)
TUS-CPS website

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

www.cancer.gov
www.cancer.gov/espanol