



Using Macros to Compute US Health Insurance Coverage Estimates for Insertion into a Web-based Table Generator

Michele Burlew, *Episystems, Inc.*,

Peter Graven, Karen Turner, Michel Boudreaux, *University of Minnesota*

Objective

Develop a system of SAS® macro programs that manage the processes of:

- 1) estimating health insurance coverage in the US population based on responses to two federal surveys
- 2) formatting the results for loading into the SHADAC (State Health Access Data Assistance Center) Data Center.

This poster describes the steps followed to design, write, and test this specific system of SAS macro programs. These steps can be adapted when designing other systems.

Understand Data Sources

Health insurance estimates are derived from two federal surveys whose results are released annually in the fall:

- the US Current Population Survey’s Annual Social and Economic Supplement (CPS), administered by the Bureau of Labor Statistics (BLS) and the Census Bureau
- the American Community Survey (ACS) administered by the Census Bureau.

Survey	Approximate Obs per Year	Data Center Estimates Start in:
CPS	200,000	1987
ACS	3,000,000	2008

Determine Variables Needed

The SAS SURVEYMEANS procedure computes the frequencies, percentages, and standard errors of the health insurance estimates for different descriptive characteristics and combinations of characteristics in the US population.

The health insurance measurements available in each survey were identified. Next, the variables that characterize the data were organized into two groups, the “Filter” variables and the “Row” variables.

Estimates are computed for subsets of the survey based on *filter variable* values. The values of *row variable* values are used to compute detail within subsets defined by the filter variables.

Organize the Processing Steps

Three major steps were identified in preparing the data for insertion into the SHADAC Data Center. Macro programs were written specifically for each step.

1. Prepare models and submit them to PROC SURVEYMEANS Produce one ODS output data set from each call to PROC SURVEYMEANS.
2. Combine all the output data sets from PROC SURVEYMEANS and post-process the data.
3. Create a CSV file from the output data set created in Step 2 to load into the web-based system.

Determine computing environment

Select variables

Recode variables to create row and filter variables

Run frequencies to validate recodes

Run SURVEYMEANS models

Save SURVEYMEANS output in data sets and PDF files

Save SAS logs

Step 1:
Analyses

Concatenate ODS output data sets from PROC SURVEYMEANS

Post-process recoded variables and remove duplicate models

Validate that all models are present

Step 2:
Combine

Create single-year CSV from combined data set

Calculate multi-year statistics from output

Create multi-year CSV

Step 3:
CSVs

Determine Analyses Needed

Models that contained valid combinations of the filter and row variables were determined. The code submits each model twice: once to analyze the US population and second to analyze each state using a BY STATE statement.

Model Category	Number of Unique Models Constructed	Total Models Analyzed Nationally and By State
Overall	1	52
Row Variables	17	884
Filter Variables x Row Variables	81	4,212
Filter Variables x Filter Variables x Row Variables	92	4,784
**Total	191	9,932

This PROC SURVEYMEANS step analyzes 2009 CPS data by state. It analyzes the subpopulations defined by the two filter variables: AGEGRP3A and FPL2CATA and the one row variable AGEROW8A. Macro code generates the step by looping through lists of filter and row variables and their acronyms to construct the statements. Acronyms in the PDF filename and ODS output data set name identify the analysis.

```
ods pdf file="p:\data center\cps\results\2009\
02 surveymeans\cps_sm_FFR_2009_20101111_abcl.pdf";
ods output domain=domainst_Fa3a_Fp2a_Rar8a;

proc surveymeans data=work.cps2009 stacking
                    mean nobS stderr sum;
    title "Processing CPS 2009";
    title2 "Filter AGEGRP3A X Filter FPL2CATA X Row
            AGEROW8A by State Model";

    by state;
    domain AGEGRP3A*FPL2CATA*AGEROW8A
            AGEGRP3A*FPL2CATA*AGEROW8A_I;
    cluster h_seq;
    strata geocode;
    var notcovered covered private empbased individual
        public medicaid medicare military;
    weight marsupwt;
    label AGEGRP3A= ' ' FPL2CATA= ' ' AGEROW8A= ' ' ;
run;
```

Develop Code in a Modular Style

Since this system of macros will analyze data annually following the release of the surveys, the code was written in a modular style to make it easier to handle changes in variable names and values that can occur over time and with changes in policy guidelines.

Macro program SEX_RECODE codes the row variable SEX in the CPS. (Variable SEX_I is needed to complete the model specification.)

```
%macro sex_recode;
    %if &cpsyear ge 1988 and &cpsyear ne 1995
        %then %let sexvar=a_sex;
    %else %if &cpsyear=1995 %then %let sexvar=pesex;
    attrib sex label='Sex'
            sex_i label="SEX not missing";
    sex=&sexvar;
    if sex=. then sex=-1;
    sex_i=(sex not in (.,-1));
%mend sex_recode;
```

Modify Code in Steps of Increasing Complexity

Because the process of producing a file that can be loaded into the web-based system involves many steps, the macros were developed at first with minimal options. As experience was gained and output reviewed, more options were added.

```
%macro docpstables(startyear=,endyear=,cpsdataset=,
                    bystate=,dofreqs=yes) / minoperator;

%macro docpstables(startyear=,endyear=,cpsdataset=,
                    bystate=,dofreqs=yes,
                    dooverall=yes,dor=yes,dofr=yes,
                    dofr=yes,initS=) / minoperator;

%macro docpstables(startyear=,endyear=,cpsdataset=,
                    bystate=,dofreqs=yes,
                    dooverall=yes,dor=yes,dofr=yes,
                    dofr=yes,
                    filterlist=all,rowlist=all,
                    initS=) / minoperator;
```

Write macros to execute with minimal intervention

The SURVEYMEANS analyses take several days to complete. Manually submitting nearly 200 models per year and survey is time-consuming and error-prone. Macro code manages the model submissions. Loops iterated through the lists of filter variables and row variables to build the SURVEYMEANS steps and save the results.

These statements are from the section of the CPS processing program that constructs the Filter x Filter x Row models.

```
. . .
%do f=1 %to &nfiltervars;
    %do g=%eval(&f+1) %to &nfiltervars;
        %do r=1 %to &nrowvars;
            %let ffilter=%upcase(%scan(&filtervars,&f));
            %let gfilter=%upcase(%scan(&filtervars,&g));
            %let rowvar=%scan(&rowvars,&r);
. . .
    %let dsname=
F%scan(&filtercodes,&f)_F%scan(&filtercodes,&g)_R%scan(&rowcodes,&r);
    %let domainstmt=&ffilter*&gfilter*&rowvar
&ffilter*&gfilter*%scan(&rowmissvars,&r);
. . .
```

Devise Rules for Naming Files and Folders

Since the processing produces many output data sets and output files, the code uses pre-defined acronyms and survey identification to name files and output. Folders are defined with specific names that the programs expect.

Save Logs and Output for Review

Because of the complexity and duration of the processing and the need to archive the results, the code saves SAS logs and procedure output in external files. PROC PRINTTO directs the SAS log to an external file. The ODS listing destination is closed and output is sent solely to a PDF file.

Screen Shots of SHADAC Data Center (www.shadac.org/datacenter)

The first screen that a user sees presents options to select the survey type, the geography, the year, and any age or poverty filter selections.

Also from SHADAC:

SHARE Supporting research on Affordable Care Act implementation at the state level.

SHAP Providing technical assistance to State Health Access Program grantees.

Home > Data Center >

Tables

Choose Your Data & Filters

Data Source: CPS
Population Results
☒ Show Count
☒ Show Percentage
☒ Show Standard Error

States: ☒ United States, ☐ Alabama, ☐ Alaska, ☐ Arizona, ☐ Arkansas, ☐ California, ☐ Colorado, ☐ Connecticut, ☐ Delaware, ☐ District of Columbia, ☐ Florida, ☐ Georgia, ☐

Year(s): Single Year Estimate, 2009

Age Range: All, 0-17, 0-18, 0-64, 18-64, 19-64, 65+

Poverty Status: All

Continue

The second screen presents options to select the detail for the tables that contain the data specified in the first screen. One or more selections of the row variables can be made.

Also from SHADAC:

SHARE Supporting research on Affordable Care Act implementation at the state level.

SHAP Providing technical assistance to State Health Access Program grantees.

Home > Data Center >

Tables

Choose Your Data & Filters

Choose Your Tables

To obtain overall estimates based on your selected data source and filters, do not select any tables in this step. Simply proceed by clicking "create table".

Table Options

☐ Race and Ethnicity
☒ Age (detailed & summary)
☐ Poverty Level (detailed)
☒ Family Income
☐ Sex
☒ Marital Status - Individual
☐ Marital Status - Family
☐ Marital/Child Status - Family
☐ Work Status - Individual
☒ Work Status - Family
☒ Education - Individual
☐ Education - Family
☐ Health Status

Create Table

The third screen shows the tables of results. The output can be saved to a CSV or PDF file.

Home > Data Center >

Tables

Choose Your Data & Filters

Choose Your Tables

Your Results

Export Tables: Excel (.csv file) PDF Save: Search Settings (Login Required)

Health Insurance Coverage Estimates, CPS, 0-64 Years, All Poverty Levels, United States: Calendar Year 2009

	Population Uninsured			Insured			Private Coverage			Government Coverage		
	Total			Total			Employer			Individual		
	Count	%	SE	Count	%	SE	Count	%	SE	Count	%	SE
Family Income												
\$0 - \$24,999	57,589	20,306	35.3	37,283	64.7	0.39	14,400	25.0	0.37	10,004	17.4	0.32
\$25,000 - \$49,999	59,713	14,049	23.5	45,663	76.5	0.35	33,608	56.3	0.45	30,083	50.4	0.45
\$50,000 - \$74,999	48,377	7,419	15.3	40,959	84.7	0.35	36,526	75.5	0.44	33,796	69.9	0.47
\$75,000+	98,985	8,224	8.2	91,764	91.8	0.19	87,597	87.6	0.24	82,660	82.7	0.28
Total	265,667	49,998	18.8	215,669	81.2	0.16	172,131	64.8	0.21	156,543	58.9	0.22
Education - Individual												
< High school	23,521	9,999	42.5	13,521	57.5	0.55	7,923	33.7	0.50	7,024	29.9	0.48
High school	57,115	16,101	28.2	41,015	71.8	0.32	33,072	57.9	0.35	30,257	53.0	0.35
Some college	56,146	10,970	19.5	45,176	80.5	0.28	39,591	70.5	0.32	35,117	62.5	0.35
College or more	53,846	5,415	10.1	48,431	89.9	0.22	46,258	85.9	0.25	42,252	78.5	0.30
Total	190,627	42,485	22.3	148,142	77.7	0.18	126,843	66.5	0.21	114,651	60.1	0.22
Filtered Total	265,667	49,998	18.8	215,669	81.2	0.16	172,131	64.8	0.21	156,543	58.9	0.22

CPS

Source: Current Population Survey Annual Social and Economic Supplement (CPS), 2010. Counts are presented in thousands.

Definitions: % = Percent; SE = Standard Error as a percent

Insurance coverage note: Listed Medicaid rates include SCHIP and state-specific public programs. Those who report multiple types of insurance coverage (i.e., private, public, or military) are included in each category; therefore, row totals may add to more than 100 percent. CPS methodology note: The CPS asks respondents about their health insurance coverage at any point in the last calendar year.

For technical documentation on the CPS please visit: <http://www.shadac.org/content/cps-info-and-resources>

Data revision: 09/30/2010

Acknowledgements

The SHADAC Data Center is funded by a grant from the Robert Wood Johnson Foundation. Data processing to prepare the estimates for SHADAC's Data Center is conducted at the Minnesota Supercomputing Institute for Advanced Computational Research. (www.msi.umn.edu)

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Contact Information

For SAS code questions, contact:
Michele Burlew
email: mmburlew@comcast.net

For SHADAC Data Center questions, contact:
SHADAC
2221 University Ave SE
Minneapolis, MN 55414
Phone: 612-624-4802 Fax: 612-624-1493
email: shadac@umn.edu