

# Analyzing MEPS-HC Data with SAS® 9.4M6 Pradip K. Muhuri, PhD April 29, 2020

# SAS® Programming Basics and Complex Survey Data Analysis



- SAS programs typically include any combination of the following:
  - DATA Step
  - PROC Step
  - Global Statements (e.g., <u>LIBNAME</u>, <u>OPTIONS</u>, and <u>TITLE/FOOTNOTE</u>)
  - Macro Variables, Macros, and Macro Functions
- Base Procedures in SAS (Examples)
  - PROC FREQ, PROC MEANS, PROC SUMMARY, PROC SORT, PROC DATASETS, PROC FORMAT, PROC PRINT, and PROC PRINTTO
- Complex Survey Procedures in SAS (Examples)
  - ► PROC SURVEYMEANS
  - PROC SURVEYFREQ
- Output Delivery Systems (ODS)
  - Controlling PROC output with ODS select/exclude
  - Saving results to a SAS data set
- Interface: SAS Windowing Environment, <u>JupyterLab</u>
- Resources for MEPS/SAS programs, code explanations, and references
  - (Primary): <a href="https://github.com/HHS-AHRQ/MEPS">https://github.com/HHS-AHRQ/MEPS</a>
  - (Supplementary): https://github.com/pkmedu/AnalyzeMEPS

### MEPS-HC SAS Transport Files on the Web



- SAS Transport (MEPS) Files
  - best overall format
  - machine-independent (data files can be moved between computers running different operating systems).
  - can be directly imported into SAS, SPSS, BMDP, and STATA, etc.

How to create a transport file for one or more data sets

# Working with MEPS-HC SAS Transport Files from the Web

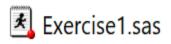


- Objectives Download\_Data\_from\_MEPS\_Site\_rev.sas
  - Run SAS macro to automate the
    - downloading of any number of SAS Transport files from the MEPS website
    - unzipping the files
    - converting them into SAS data sets
- Macro that wraps the following:
  - ▶ PROC HTTP for Data Downloads
  - How to Unzip/Read Data Files in SAS
  - ► CALL SYMPUTX Routine
  - PROC COPY XPORT Engine

Here is a non-macro SAS program that handles one file at a time (AnalyzeMEPS Repo on GitHub).



Objective



- Generate the following estimates
  - mean health care expenses per person
  - mean health expenses per person with an expense (overall, and by age group)
- Data and Analysis
  - ▶ Use the 2017 MEPS Full-Year Consolidated File
  - Create a subpopulation variable for DOMAIN analysis
  - Run PROC FREQ for data checks
  - ► Run PROC SURVEYMEANS for complex survey estimates



- Objective Exercise2.sas
  - Estimate the following parameters
    - Annual mean number of purchases of opioids (i.e., Narcotic analgesics or Narcotic analgesic combos) per person with one or more purchases of opioids
    - Annual mean total, out-of-pocket, and third-party payer expenses for purchases of opioids per person with one or more purchases of opioids
- Data and Analysis
  - Aggregate 2017 MEPS prescribed medicines data at the personlevel
  - Merge aggregated prescribed medicine data with full-year personlevel data for the same year
  - ► Create a subpopulation variable for DOMAIN analysis
  - Run PROC FREQ for data checks
  - Run PROC SURVEYMEANS for complex survey estimates



- Objective Exercise3.sas
  - Estimate annual mean out-of-pocket health care expenses for individuals who were aged 26-30 years with high income and uninsured for the whole year (averaged over two years)
- Data and Analysis
  - Combine data from 2016 and 2017 MEPS Full-Year Consolidated Files
  - Create a subpopulation variable for DOMAIN analysis
  - Create a new variable (i.e., pooled sample weight) for pooled data set (2016 and 2017 combined)
  - Run PROC FREQ and PROC MEANS for data checks
  - Run PROC SURVEYMEANS for complex survey estimates



- Objective Exercise4.sas
  - Estimate the percentage distribution of insurance status (in the second year) of individuals who were aged 26-30 with high income and uninsured for the whole (first) year (averaged over three panels)
- Data and Analysis
  - Combine data from MEPS Longitudinal Files (Panels 19, 20, and 21)
  - Create a subpopulation variable for DOMAIN analysis
  - Create a new variable (i.e., pooled sample weight) for pooled data set (3 panels combined)
  - Run PROC FREQ and PROC MEANS for data checks
  - Run PROC SURVEYMEANS for complex survey estimates

### PROC SURVEYFREQ vs. PROC SURVEYMEANS



- PROC SURVEYFREQ and PROC SURVEYMEANS with a CLASS statement produce identical results (percentage vs. proportion).
- PROC SURVEYFREQ treats the variable in the TABLES statement as categorical and estimate the percentage in each category or level.
- The CLASS statement in PROC SURVEYMEANS treats the variable in the VAR statement as categorical and estimate the proportion in each category or level.

### PROC SURVEYMEANS vs. PROC MEANS



 PROC SURVEYMEANS and PROC MEANS with a WEIGHT statement produce the same results for the mean, not the confidence interval of the mean.

# PROC SURVEYMEANS Output Objects



- Use ODS TRACE statements that produces the record containing at least the following items for PROC SURVEMEANS
  - Name
  - Label
  - ▶ Template
  - Path

Controlling PROC output with ODS select/exclude

See the SAS program and SAS Log here (AnalyzeMEPS

Repo on GitHub)

#### **ODS SELECT/EXCLUDE**



- SAS Procedures like PROC SURVEYMEANS produces lot of output
  - Summary
  - Statistics
  - Quantiles
  - Domain
  - DomainQuantiles
- ODS SELECT or EXCLUDE statement tells SAS
  - What output to print
  - What output not to print

# SAVING PROC SURVEYMEANS Output



- Estimate descriptive statistics including Q1, Median, and Q3
- Save specific output tables to separate SAS data sets for the
  - overall population
  - DOMAIN of interest
  - multiple table names and data sets names allowed in the following statement

ODS OUTPUT <table-name>= <data-set-name>;

### Comparing Domain Means with PROC SURVEYMEANS



- Pairwise comparisons of the estimate among domain levels
- Bonferroni multiple comparison adjustment for the p-values for testing differences in the analysis variable among domain levels

### Sample SAS Macro – Generate Estimates from 2009 to 2019



```
%macro runit (byvar, fmt, first=, last=);
%do yr=&first %to &last;
title "MEPS, 20%sysfunc(putn(&yr,z2.))";
ods graphics off;
ods exclude statistics;
      proc surveymeans data=new.summary_person_%sysfunc(putn(&yr,z2.));
      stratum varstr;
      cluster varpsu;
      weight perwtf;
      var hd;
      domain age_18p('1');
        format &byvar &fmt;
           ods output domain=domain| %sysfunc(putn(&yr,z2.));
      run;
%end;
%mend runit:
%runit(age_grp, ageF., first=09, last=15)
```

### **Questions**



Any Questions?
Thanks!!!