

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Analyzing MEPS-HC Data with SAS® 9.4 M6

Pradip K. Muhuri, PhD

March 31, 2022

SAS® Programming Basics and Complex Survey Data Analysis

- SAS programs typically include any combination of the following:
  - DATA Step (Reading external data files, Combining SAS data sets, Manipulating data, Using data set options, PUT Function, CATS Function, STRIP Function,
  - LENGTH Statement
  - Summing Numeric Variables Using By Group (PROC SUMMARY)
  - PROC Step (Base Procedures-1 including PROC PRINT and PROC FREQ, Base Procedures-2)
  - Global Statements (e.g., LIBNAME, OPTIONS, and TITLE/FOOTNOTE)
  - Macro Variables, Macros, and Macro Functions
  - Routing Log and Output to External Files using PROC PRINTTO Statement
- Complex Survey Procedures in SAS (Examples)
  - PROC SURVEYMEANS , PROC SURVEYFREQ
  - PROC SURVEYREG, PROC SURVEYLOGISTC
- Output Delivery Systems (ODS)
  - Controlling PROC output with ODS select/exclude
  - Saving results to a SAS data set
- Interface
  - SAS Windowing Environment
  - JupyterLab
- Resources for MEPS/SAS programs, code explanations, and references
  - (Primary): https://github.com/HHS-AHRQ/MEPS-workshop/tree/master/sas_exercises
  - (Supplementary): https://github.com/pkmedu/AnalyzeMEPS

SAS PROCs for Complex Survey Data Analysis

- Use PROC SURVEYMEANS to obtain weighted descriptive statistics (e.g., means, medians, proportions) and produce graphs.
- Use PROC SURVEYFREQ to obtain weighted one-way or multi-way crosstabulations (e.g., percentages) and produce graphs.
- Use PROC SURVEYREG to run weighted OLS regressions (not covered in this session)
- Use PROC SURVEYLOGISTIC to run weighted logistic, ordinal, multinomial and probit regressions.
- Use STRATA, CLUSTER, and WEIGHT statements (required) for variance estimation when running the above complex survey procs.
- Use the Use the proper variance structure when making estimates from MEPS data pooled over multiple years. Read MEPS guidelines on pooling data.
- Use the DOMAIN statement to define domains of interest for all complex survey PROCs. Especially for domain analyses via PROC SURVEYMEANS (with SAS 9.4 M5 or later versions), you can write a DOMAIN statement based on a single domain level from one or more domain variables as follows:
  - DOMAIN agelast('65+')*Mental_disorders('1');
- Do not filter your data set to estimate domain statistics. Variance estimates for domain statistics estimated via the WHERE or the BY statement might not be valid. (SAS® Documentation)