

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Analyzing MEPS-HC Data with SAS® 9.4 M6

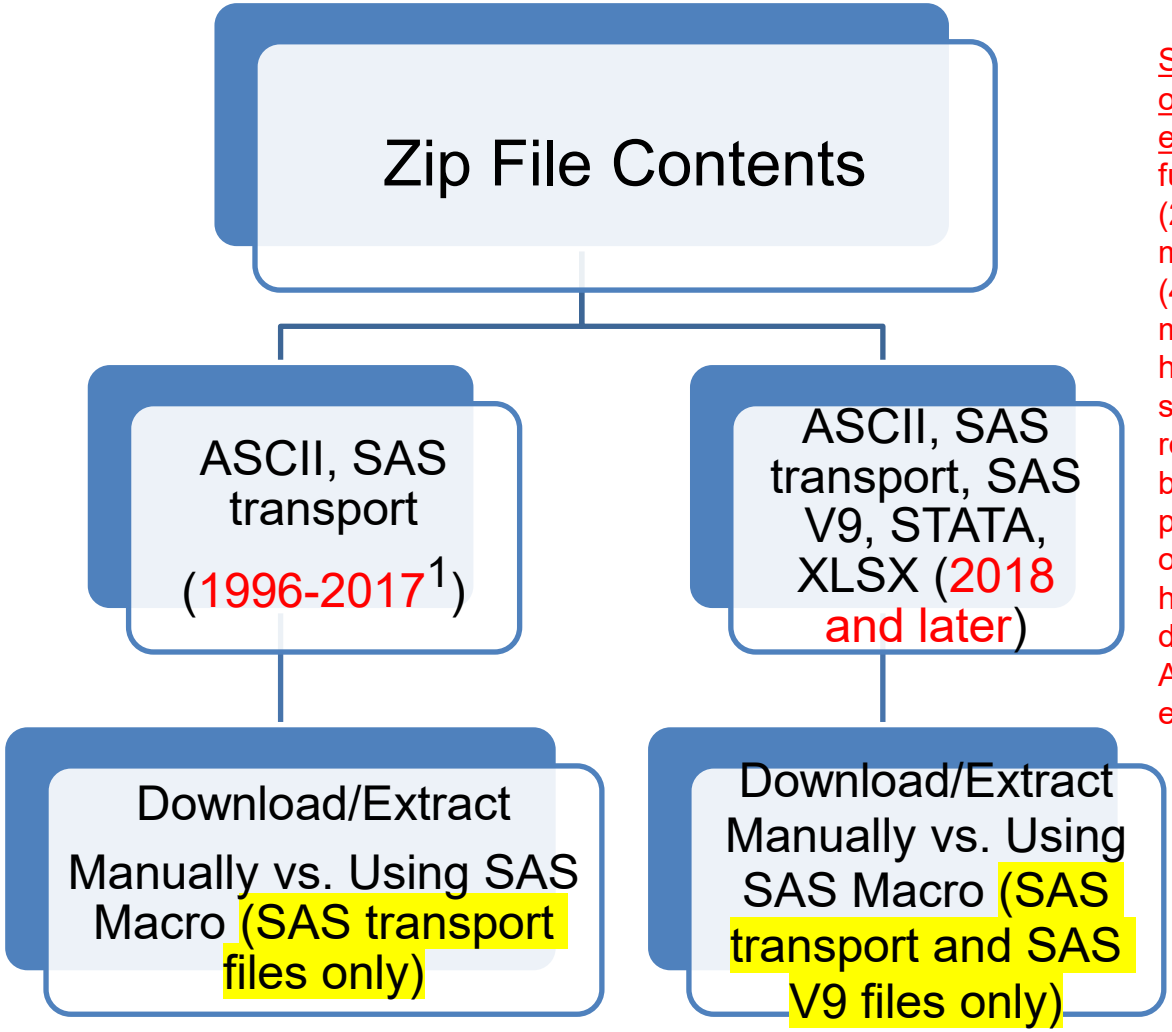
Pradip K. Muhuri, PhD

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

MEPS Data File Formats and Ways to Download/Extract



Some common types of MEPS data files for each data year: (1) full year consolidated, (2) longitudinal, (3) medical conditions, (4) prescribed medicines, (5) hospital inpatient stays, (6) emergency room visits, (7) office-based medical provider visits, (8) outpatient visits, (9) home health, (10) dental visits, and (11) Appendix to MEPS events.

[1]One exception is the Full Year Consolidated File for 2017, which is available in three additional formats (i.e., SAS V9, STATA, and XLSX).

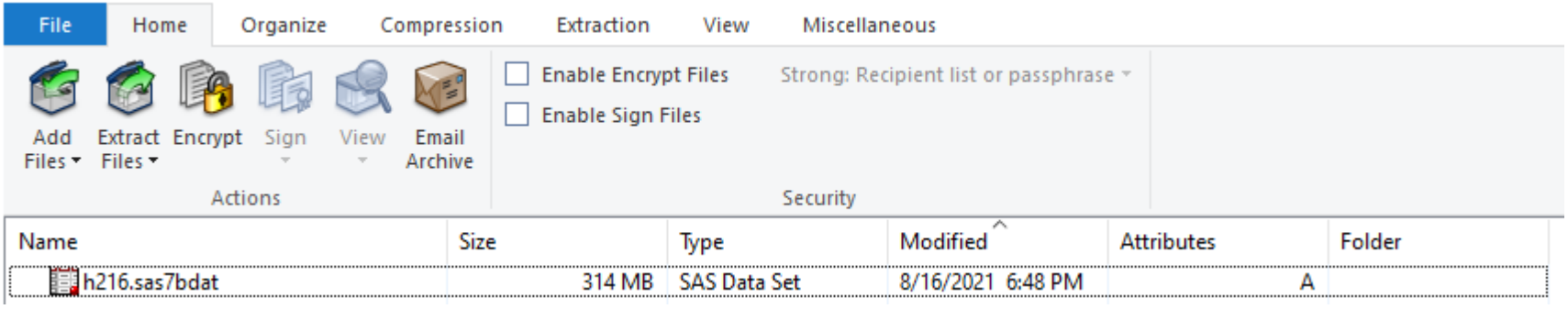
Download/Extract Files Manually (Example: MEPS HC-216 SAS V9 Files)

MEPS HC-216: 2019 Full Year Consolidated Data File

https://meps.ahrq.gov/data_stats/download_data_files_detail.jsp?cboPufNumber=HC-216. (1) Go to the bottom of the above URL page and see data files of multiple formats. (2) Click on ZIP for the third Data File, SAS V9 format.

Data	File type
Data File, ASCII format	ZIP (11 MB) / EXE (12 MB)
Data File, SAS transport format	ZIP (11 MB) / EXE (12 MB)
Data File, SAS V9 format	ZIP (13 MB)
Data File, Stata format	ZIP (13 MB)
Data File, XLSX format	ZIP (84 MB)

(3) Wait until the file-download is completed and SecureZip menu items are available (screenshot below). (4) Click on Extract Files and follow the instruction to extract/save the single .sas7bdat file to C:\MEPS Data.



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 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)

PROC SURVEYMEANS
(Means and Proportions)

```
Data work.h209;
  set pufmeps.h209;
  if INSCOV18=3 then Unsinsured_d=1; else Unsinsured_d=0;
  if INSCOV18=3 then Unsinsured_s='Uninsured'; else Unsinsured_s='Insured';
  Label Unsinsured_d = 'Whether uninsured (1/0 variable)'
        Unsinsured_s = 'Whether uninsured (String variable)';
run;
ods graphics off;
PROC SURVEYMEANS DATA=work.h209 ;
   VAR totexp18  Unsinsured_d Unsinsured_s INSCOV18;
   STRATUM VARSTR;
   CLUSTER VARPSU;
   WEIGHT PERWT18F;
   CLASS INSCOV18;
   FORMAT INSCOV18 INSCOV18_fmt.;
RUN;
```

You can request means, and proportions from PROC SURVEYMEANS. Multiple variables (mix of character, categorical, and numeric variables) are allowed in the VAR statement.

The CLASS statement in PROC SURVEYMEANS treats the variable (INSCOV18) in the VAR statement as categorical and estimate the proportion in each category or level.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: ProcSurveyMeansPercentiles.SAS (https://github.com/pkmedu/AnalyzeMEPS).

PROC SURVEYMEANS (Means and Proportions) - Output

The SURVEYMEANS Procedure

Data Summary

Number of Strata	117
Number of Clusters	257
Number of Observations	30461
Number of Observations Used	29415
Number of Obs with Nonpositive Weights	1046
Sum of Weights	326327888

ODS Tables produced by PROC SURVEYMEANS

- SUMMARY
- STATISTICS (estimate over the entire population)

Class Level Information

Variable	Label	Levels	Values
Unsinsured_s	Whether uninsured (String variable)	2	Insured Uninsured
INSCOV18	HEALTH INSURANCE COVERAGE INDICATOR 2018	3	Any Private Public Only Uninsured

Statistics

Variable	Level	Label	N	Mean	Std Error of Mean	95% CL for Mean	
TOTEXP18		TOTAL HEALTH CARE EXP 18	29415	6063.134458	128.011022	5810.04979	6316.21912
Unsinsured_d		Whether uninsured (1/0 variable)	29415	0.065089	0.002930	0.05930	0.07088
Unsinsured_s	Insured	Whether uninsured (String variable)	27248	0.934911	0.002930	0.92912	0.94070
	Uninsured	Whether uninsured (String variable)	2167	0.065089	0.002930	0.05930	0.07088
INSCOV18	Any Private	HEALTH INSURANCE COVERAGE INDICATOR 2018	17633	0.675677	0.008286	0.65929	0.69206
	Public Only	HEALTH INSURANCE COVERAGE INDICATOR 2018	9615	0.259234	0.006816	0.24576	0.27271
	Uninsured	HEALTH INSURANCE COVERAGE INDICATOR 2018	2167	0.065089	0.002930	0.05930	0.07088

PROC SURVEYFREQ

```
PROC SURVEYFREQ DATA=pufmeps.h209;
     TABLES INSCOV18;
     STRATUM VARSTR;
     CLUSTER VARPSU;
     WEIGHT PERWT18F;
     FORMAT INSCOV18 INSCOV18_fmt.;
RUN;
```

PROC SURVEYFREQ treats the variable in the TABLES statement as categorical and estimate the percentage in each category or level.

For a categorical variable, PROC SURVEYFREQ produces identical results from PROC SURVEYMEANS with CLASS statement (percentage vs. proportion).

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: ProcSurvefreq.SAS (https://github.com/pkmedu/AnalyzeMEPS).

PROC SURVEYFREQ

The SURVEYFREQ Procedure

Data Summary

Number of Strata	117
Number of Clusters	257
Number of Observations	30461
Number of Observations Used	29415
Number of Obs with Nonpositive Weights	1046
Sum of Weights	326327888

HEALTH INSURANCE COVERAGE INDICATOR 2018

INSCOV18	Frequency	Weighted Frequency	Std Err of Wgt Freq	Percent	Std Err of Percent
Any Private	17633	220492239	6158571	67.5677	0.8286
Public Only	9615	84595319	2640781	25.9234	0.6816
Uninsured	2167	21240330	959379	6.5089	0.2930
Total	29415	326327888	7295775	100.0000	

ODS Tables produced by PROC SURVEYFREQ

- SUMMARY
- ONEWAY

PROC SURVEYMEANS

```
PROC SURVEYMEANS DATA=pufmeps.h209;
VAR totexp18;
STRATUM VARSTR;
CLUSTER VARPSU;
WEIGHT PERWT18F;
DOMAIN POVCAT18;
FORMAT POVCAT18 povcat_fmt.;
RUN;
```

- When you request domain analysis from PROC SURVEYMEANS, you still get the estimates over the entire population.
- Key estimates: population totals and means of total annual health care spending for overall and 4 income domains.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: ProcSurveyMeansPercentiles.SAS (https://github.com/pkmedu/AnalyzeMEPS).

PROC SURVEYMEANS - Output

The SAS System
The SURVEYMEANS Procedure

Data Summary

Number of Strata	117
Number of Clusters	257
Number of Observations	30461
Number of Observations Used	29415
Number of Obs with Nonpositive Weights	1046
Sum of Weights	326327888

ODS Tables produced by the PROC SURVEYMEANS step which includes the DOMAIN statement

- SUMMARY
- STATISTICS (estimate over the entire population)
- DOMAIN

Statistics

Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
TOTEXP18	TOTAL HEALTH CARE EXP 18	29415	6063.134458	128.011022	5810.04979	6316.21912

The SAS System
The SURVEYMEANS Procedure

Statistics for POVCAT18 Domains

POVCAT18	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
Poor	TOTEXP18	TOTAL HEALTH CARE EXP 18	5186	6387.431789	325.672968	5743.55884	7031.30474
Near Poor/Low Income	TOTEXP18	TOTAL HEALTH CARE EXP 18	5892	5749.822666	220.769022	5313.35045	6186.29488
Middle Income	TOTEXP18	TOTAL HEALTH CARE EXP 18	8409	5593.408276	246.604199	5105.85853	6080.95803
High Income	TOTEXP18	TOTAL HEALTH CARE EXP 18	9928	6424.070915	209.836076	6009.21372	6838.92811

Working with ODS Tables for Selected Complex Survey PROCs

- ODS Table Names for PROC SURVEYMEANS (Examples)
  - SUMMARY
  - STATISTICS
  - DOMAIN
  - QUANTILES
  - DOMAINQUANTILES
  - DOMAINDIFFS

  ODS Table Names for PROC SURVEYFREQ

  ODS Table Names for PROC SURVEYLOGISTIC

  See HowToListODSTableNamesByTracingOn.sas (https://github.com/pkmedu/AnalyzeMEPS)

- Use ODS SELECT/ ODS EXCLUDE to control for output destination
  - What output to print (e.g., ODS SELECT DOMAIN)
  - What output not to print (e.g., ODS EXCLUDE ALL)

- SAS Statement in PROC SURVEYMEANS code block

```
ODS OUTPUT DOMAIN=MEAN_EST DOMAINQUANTILES=QUANT_EST;
```

Generating Reports from ODS Table-Based SAS Data Set with PROC SURVEMEANS (Part 1)

```
ODS SELECT DOMAINQUANTILES;
PROC SURVEYMEANS DATA=PUFMEPS.H209 NOBS Q1 MEDIAN Q3;
VAR TOTEXP18;
STRATUM VARSTR;
CLUSTER VARPSU;
WEIGHT PERWT18F;
DOMAIN POVCAT18;
FORMAT POVCAT18 POVCAT_FMT.;
RUN;
ODS SELECT ALL;
```

Here you request 25th , 50th, and 75th percentiles of total annual health care spending by income levels (domain analysis) from PROC SURVEYMEANS. This code block does not generate estimates over the entire population! Why? Because the ODS SELECT statement restricts the output to the "DOMAINQUANTILES" table.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: `ProcSurveyMeansPercentiles.SAS` (https://github.com/pkmedu/AnalyzeMEPS).

Generating Reports from ODS Table-Based SAS Data Set with PROC SURVEMEANS (Part 1) - Output

The SURVEYMEANS Procedure

Quantiles for POVCAT18 Domains

POVCAT18	Variable	Percentile	Estimate	Std Error	95% Confidence Limits	
Poor	TOTEXP18	25 Q1	113.411301	13.291096	87.13409	139.68851
		50 Median	792.584635	63.304044	667.42914	917.74013
		75 Q3	4641.089476	371.107961	3907.38910	5374.78985
Near Poor/Low Income	TOTEXP18	25 Q1	158.365316	13.634919	131.40835	185.32228
		50 Median	893.057345	50.119031	793.96933	992.14536
		75 Q3	4238.327426	211.254428	3820.66607	4655.98878
Middle Income	TOTEXP18	25 Q1	228.505336	18.096932	192.72673	264.28394
		50 Median	1120.760529	57.196205	1007.68056	1233.84050
		75 Q3	4285.295502	173.772478	3941.73798	4628.85303
High Income	TOTEXP18	25 Q1	476.705475	18.503398	440.12326	513.28769
		50 Median	1751.713707	56.789438	1639.43794	1863.98948
		75 Q3	5809.573428	163.337074	5486.64726	6132.49960

Note: Due to the ODS SELECT DOMAINQUANTILES; statement, the output is limited to DOMAINQUANTILES, and there is no output for SUMMARY and STATISTICS here.

PROC SURVEYMEANS (Subpopulation Analysis)

```
ODS SELECT DOMAIN;
PROC SURVEYMEANS DATA=pufmeps.h209 ;
     VAR totexp18;
     STRATUM VARSTR;
     CLUSTER VARPSU;
     WEIGHT PERWT18F;
DOMAIN POVCAT18('High Income')*INSCOV18('ANY PRIVATE');
FORMAT POVCAT18 povcat_fmt.INSCOV18 inscov_fmt.;
RUN;
ODS SELECT ALL;
```

The ODS SELECT statement restricts the output to "DOMAIN".

This quoted formatted value syntax in the DOMAIN statement controls only the display of domain analysis results; it does not subset the data set, change the degrees of freedom, or otherwise affect the variance.

This statement resets to all output.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: ProcSurveyMeansPercentiles.SAS (https://github.com/pkmedu/AnalyzeMEPS).

PROC SURVEYMEANS (Subpopulation Analysis) - Ourput

The SURVEYMEANS Procedure

Statistics for POVCAT18*INSCOV18 Domains

POVCAT18	INSCOV18	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean	
High Income	ANY PRIVATE	TOTEXP18	TOTAL HEALTH CARE EXP 18	8612	6303.286899	218.611344	5871.08053	6735.49327

Note: Due to the ODS SELECT DOMAIN; statement before PROC SURVEYMEANS, the output is limited to DOMAIN, and there is no output for SUMMARY and STATISTICS here.

Generating Reports from ODS Table-Based Output Data Set in PROC SURVEMEANS (Part 1)

```
ODS EXCLUDE ALL;
PROC SURVEYMEANS DATA=PUFMEPS.H209 NOBS MEAN STDERR Q1 MEDIAN Q3;
    VAR TOTEXP18;
    STRATUM VARSTR;
    CLUSTER VARPSU;
    WEIGHT PERWT18F;
    DOMAIN POVCAT18;
    FORMAT POVCAT18 POVCAT_FMT.;
ODS OUTPUT DOMAIN=MEAN_EST DOMAINQUANTILES=MED_EST;
RUN;
ODS SELECT ALL; *RESTORE THE DEFAULT PRINTED OUTPUT;
TITLE 'MEAN ESTIMATES';
PROC PRINT DATA=WORK.MEAN_EST;
  VAR POVCAT18   N  MEAN  STDERR;
RUN;
TITLE 'PERCENTILE ESTIMATES';
PROC PRINT DATA=WORK.MED_EST;
  VAR  POVCAT18 PERCENTILE ESTIMATE STDERR  LOWERCL  UPPERCL;
RUN;
TITLE;
```

Here you request estimate of mean, 25th , 50th, and 75th percentiles of annual total health care spending by income levels from PROC SURVEYMEANS.

- ODS EXCLUDE ALL suspends the SAS output.
- ODS OUTPUT statement creates two SAS data sets based on ODS Tables (DOMAIN and DOMAINQUANTILES).
- PROC PRINT provides listing of estimates from the SAS data sets.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: ProcSurveymeansPercentiles.SAS (https://github.com/pkmedu/AnalyzeMEPS).

.

Generating Reports from ODS Table-Based SAS Data Set in PROC SURVEMEANS – Output (Part 1)

Mean estimates

Obs	POVCAT18	N	Mean	StdErr
1	Poor	5186	6387.431789	325.672968
2	Near Poor/Low Income	5892	5749.822666	220.769022
3	Middle Income	8409	5593.408276	246.604199
4	High Income	9928	6424.070915	209.836076

Quartile estimates

Obs	POVCAT18	Percentile	Estimate	StdErr	LowerCL	UpperCL
1	Poor	25	113.411301	13.291096	87.13409	139.68851
2	Poor	50	792.584635	63.304044	667.42914	917.74013
3	Poor	75	4641.089476	371.107961	3907.38910	5374.78985
4	Near Poor/Low Income	25	158.365316	13.634919	131.40835	185.32228
5	Near Poor/Low Income	50	893.057345	50.119031	793.96933	992.14536
6	Near Poor/Low Income	75	4238.327426	211.254428	3820.66607	4655.98878
7	Middle Income	25	228.505336	18.096932	192.72673	264.28394
8	Middle Income	50	1120.760529	57.196205	1007.68056	1233.84050
9	Middle Income	75	4285.295502	173.772478	3941.73798	4628.85303
10	High Income	25	476.705475	18.503398	440.12326	513.28769
11	High Income	50	1751.713707	56.789438	1639.43794	1863.98948
12	High Income	75	5809.573428	163.337074	5486.64726	6132.49960

Generating Reports from ODS Table-Based SAS Data Set in PROC SURVEMEANS (Part 2)

```
ODS EXCLUDE ALL;
PROC SURVEYMEANS DATA=PUFMEPS.H209 NOBS SUMWGT SUM MEAN STDERR;
    VAR TOTEXP18 TOTSLF18 OBTOTV18;
    STRATUM VARSTR;
    CLUSTER VARPSU;
    WEIGHT PERWT18F;
    DOMAIN POVCAT18;
    FORMAT POVCAT18 POVCAT_FMT.;
    ODS OUTPUT DOMAIN=WORK.DOMAIN_RESULTS;
RUN;
ODS SELECT ALL;
PROC PRINT DATA=WORK.DOMAIN_RESULTS (DROP=DOMAINLABEL)
                NOOBS LABEL BLANKLINE=3 SPLIT='*';
LABEL POVCAT18 = 'FAMILY INCOME'
      SUMWGT = 'SUM OF*THE WEIGHTS'
      SUM = 'ESTIMATED*POPULATION*TOTAL*VARIABLE*Y';
VAR POVCAT18 VARNAME N SUMWGT SUM STDDEV MEAN STDERR;
FORMAT N                        COMMA6.0
       SUMWGT    SUM     STDDEV COMMA17.0
       MEAN      STDERR         COMMA9.2
       POVCAT18                 POVCAT_FMT. ;
RUN;
```

Here you request estimates of annual total health care spending (TOTEXP18), out-of-pocket expenses (TOTSLF18), and office-based expenses (OBTOTV18) for four domains (family income) from PROC SURVEYMEANS.

- ODS EXCLUDE ALL suspends the SAS output.
- ODS OUTPUT statement creates a SAS data set based on ODS Table (DOMAIN)
- .
- ODS SELECT ALL resets to all output tables.
- PROC PRINT provides listing of estimates from the SAS data set (DOMAIN_RESULTS). FORMAT is applied to TOTEXP18, TOTSLF18, and OBTOTV18.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: ProcSurveymeansPercentiles.SAS (https://github.com/pkmedu/AnalyzeMEPS).

Generating Reports from ODS Table-Based SAS Data Set in PROC SURVEYMEANS – Output (Part 2)

FAMILY INCOME LEVEL	Variable Name	N	SUM OF THE WEIGHTS	ESTIMATED POPULATION TOTAL VARIABLE Y	Std Error of Sum	Mean	Std Error of Mean
Poor	TOTEXP18	5,186	38,937,040	248,707,688,886	15,377,398,894	6,387.43	325.67
Poor	TOTSLF18	5,186	38,937,040	13,827,564,732	1,007,528,022	355.13	23.61
Poor	OBTOTV18	5,186	38,937,040	248,899,861	15,615,027	6.39	0.31
Near Poor/Low Income	TOTEXP18	5,892	56,018,645	322,097,274,165	15,062,570,475	5,749.82	220.77
Near Poor/Low Income	TOTSLF18	5,892	56,018,645	27,000,931,419	1,748,703,771	482.00	29.53
Near Poor/Low Income	OBTOTV18	5,892	56,018,645	326,936,163	14,814,144	5.84	0.21
Middle Income	TOTEXP18	8,409	94,607,047	529,175,839,921	29,090,820,376	5,593.41	246.60
Middle Income	TOTSLF18	8,409	94,607,047	72,032,658,555	3,807,172,369	761.39	34.59
Middle Income	OBTOTV18	8,409	94,607,047	539,527,176	20,931,085	5.70	0.16
High Income	TOTEXP18	9,928	136,765,156	878,589,058,564	39,570,494,086	6,424.07	209.84
High Income	TOTSLF18	9,928	136,765,156	152,494,041,666	7,517,091,569	1,115.01	39.54
High Income	OBTOTV18	9,928	136,765,156	970,516,807	39,396,281	7.10	0.18

Pairwise Comparisons of Estimates with PROC SURVEYMEANS

```
ODS GRAPHICS OFF; /*SUPPRESS THE GRAPHICS */
PROC SURVEYMEANS DATA=WORK.H209;
    VAR  TOTEXP18;
    STRATUM VARSTR;
    CLUSTER VARPSU;
    WEIGHT PERWT18F;
    DOMAIN POVCAT18 / DIFFMEANS ;
    FORMAT POVCAT18 POVCAT_FMT.;
RUN;
```

Code explanation: The DIFFMEANS option requests pairwise comparison of total health care spending (TOTEXP18)among income levels.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: HowToTestSurveyMeans.SAS (https://github.com/pkmedu/AnalyzeMEPS).

Pairwise Comparisons of Estimates with PROC SURVEYMEANS – Output

The SURVEYMEANS Procedure

Data Summary

Number of Strata	117
Number of Clusters	257
Number of Observations	30461
Number of Observations Used	29415
Number of Obs with Nonpositive Weights	1046
Sum of Weights	326327888

Statistics

Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean
TOTEXP18	TOTAL HEALTH CARE EXP 18	29415	6063.134458	128.011022	5810.04979 6316.21912

The SURVEYMEANS Procedure

Statistics

Variable	N	Mean	Std Error of Mean	95% CL for Mean
TOTEXP18	29415	6063.134458	128.011022	5810.04979 6316.21912

POVCAT18	Variable	N	Mean	Std Error of Mean	95% CL for Mean
Poor	TOTEXP18	5186	6387.431789	325.672968	5743.55884 7031.30474
Near Poor/Low Income	TOTEXP18	5892	5749.822666	220.769022	5313.35045 6186.29488
Middle Income	TOTEXP18	8409	5593.408276	246.604199	5105.85853 6080.95803
High Income	TOTEXP18	9928	6424.070915	209.836076	6009.21372 6838.92811

Differences of TOTEXP18 Means for POVCAT18 **Domains**

POVCAT18	-POVCAT18	Diff Estimate	Std Error	DF	t Value	Pr > |t|
Poor	Near Poor/Low Income	637.609124	385.010481	140	1.66	0.0999
Poor	Middle Income	794.023513	411.909949	140	1.93	0.0559
Poor	High Income	-36.639126	410.046571	140	-0.09	0.9289
Near Poor/Low Income	Middle Income	156.414390	329.798619	140	0.47	0.6360
Near Poor/Low Income	High Income	-674.248250	291.591788	140	-2.31	0.0222
Middle Income	High Income	-830.662639	311.350011	140	-2.67	0.0085

Bonferroni Multiple Comparison Tests of Pairwise Differences in Estimates with PROC SURVEYMEANS

```
ODS EXCLUDE SUMMARY STATISTICS;
PROC SURVEYMEANS DATA=pufmeps.H209 ;
     VAR TOTEXP18;
     STRATUM VARSTR;
     CLUSTER VARPSU;
     WEIGHT PERWT18F;
     DOMAIN POVCAT18 / ADJUST=BON ;
    FORMAT POVCAT18 POVCAT_FMT.;
RUN;
```

Code explanation: The ODS EXCLUDE statement excludes ODS tables for SUMMARY and STATISTICS from output destinations. The ADJUST=BON option requests a Bonferroni multiple comparison adjustment for the p-values for testing the pairwise differences in total health care spending (TOTEXP18)among poverty-levels (POVCAT18). This option also invokes the DIFFMEANS option.

Five Reasons to use the ODS EXCLUDE statement.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: HowToTestSurveyMeans.SAS (https://github.com/pkmedu/AnalyzeMEPS).

Bonferroni Multiple Comparison Tests of Pairwise Differences in Estimates with PROC SURVEYMEANS - Output

The SURVEYMEANS Procedure

Statistics for POVCAT18 Domains

POVCAT18	Variable	Label	N	Mean	Std Error of Mean	95% CL for Mean
Poor	TOTEXP18	TOTAL HEALTH CARE EXP 18	5186	6387.431789	325.672968	5743.55884 7031.30474
Near Poor/Low Income	TOTEXP18	TOTAL HEALTH CARE EXP 18	5892	5749.822666	220.769022	5313.35045 6186.29488
Middle Income	TOTEXP18	TOTAL HEALTH CARE EXP 18	8409	5593.408276	246.604199	5105.85853 6080.95803
High Income	TOTEXP18	TOTAL HEALTH CARE EXP 18	9928	6424.070915	209.836076	6009.21372 6838.92811

Differences of TOTEXP18 (TOTAL HEALTH CARE EXP 18) Means for POVCAT18 Domains

POVCAT18	-POVCAT18	Diff Estimate	Std Error	DF	t Value	Pr > |t|	Adj P
Poor	Near Poor/Low Income	637.609124	385.010481	140	1.66	0.0999	0.5997
Poor	Middle Income	794.023513	411.909949	140	1.93	0.0559	0.3355
Poor	High Income	-36.639126	410.046571	140	-0.09	0.9289	1.0000
Near Poor/Low Income	Middle Income	156.414390	329.798619	140	0.47	0.6360	1.0000
Near Poor/Low Income	High Income	-674.248250	291.591788	140	-2.31	0.0222	0.1333
Middle Income	High Income	-830.662639	311.350011	140	-2.67	0.0085	0.0512

Confidence Limits for Pairwise Differences in Estimates with PROC SURVEYMEANS

```
ODS SELECT DOMAINDIFFS;
PROC SURVEYMEANS DATA=pufmeps.H209 ;
     VAR TOTEXP18;
     STRATUM VARSTR;
     CLUSTER VARPSU;
     WEIGHT PERWT18F;
     DOMAIN POVCAT18 / CLDIFF;
     FORMAT POVCAT18 POVCAT_FMT.;
 RUN;
```

Code explanation: The ODS SELECT restricts the output to "DOMAINDIFFS". CLDIFF requests t type confidence limits for each difference of domain means. If you leave this option in the code, it automatically invokes the DIFFMEANS option.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: HowToTestSurveyMeans.SAS (https://github.com/pkmedu/AnalyzeMEPS).

Confidence Limits for Pairwise Differences in Estimates with PROC SURVEYMEANS - Output

The SURVEYMEANS Procedure

Differences of TOTEXP18 Means for POVCAT18 Domains

POVCAT18	-POVCAT18	Diff Estimate	Std Error	DF	t Value	Pr > |t|	95% Confidence Limits	
Poor	Near Poor/Low Income	637.609124	385.010481	140	1.66	0.0999	-123.5773	1398.79553
Poor	Middle Income	794.023513	411.909949	140	1.93	0.0559	-20.3446	1608.39161
Poor	High Income	-36.639126	410.046571	140	-0.09	0.9289	-847.3232	774.04497
Near Poor/Low Income	Middle Income	156.414390	329.798619	140	0.47	0.6360	-495.6152	808.44398
Near Poor/Low Income	High Income	-674.248250	291.591788	140	-2.31	0.0222	-1250.7409	-97.75562
Middle Income	High Income	-830.662639	311.350011	140	-2.67	0.0085	-1446.2183	-215.10694

Note: Due to the ODS SELECT DOMAINDIFFS; statement before PROC SURVEYMEANS, the output is limited to DOMAINDIFFS, and there is no output for SUMMARY, STATISTICS and DOMAIN here.

Bonferroni Multiple Comparison Tests and Confidence Limits with PROC SURVEYMEANS

```
ODS EXCLUDE SUMMARY STATISTICS;
PROC SURVEYMEANS DATA=pufmeps.H209 ;
          VAR TOTEXP18;
          STRATUM VARSTR;
          CLUSTER VARPSU;
          WEIGHT PERWT18F;
          DOMAIN POVCAT18 / ADJUST=BON CLDIFF;
          FORMAT POVCAT18 POVCAT_FMT.;
RUN;
```

Code explanation: See previous slides for explanation of the ODS EXCLUDE statement and the ADJUST=BON and CLDIFF options.

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement, global statements (e.g., OPTIONS and ODS statements), and PROC FORMAT (not shown in this slide). See the SAS output in the next slide.

Source: HowToTestSurveyMeans.SAS (https://github.com/pkmedu/AnalyzeMEPS).

Bonferroni Multiple Comparison Tests and Confidence Limits with PROC SURVEYMEANS - Output

The SURVEYMEANS Procedure

Differences of TOTEXP18 Means for POVCAT18 Domains

POVCAT18	-POVCAT18	Diff Estimate	Std Error	DF	t Value	Pr > |t|	Adj P
Poor	Near Poor/Low Income	637.609124	385.010481	140	1.66	0.0999	0.5997
Poor	Middle Income	794.023513	411.909949	140	1.93	0.0559	0.3355
Poor	High Income	-36.639126	410.046571	140	-0.09	0.9289	1.0000
Near Poor/Low Income	Middle Income	156.414390	329.798619	140	0.47	0.6360	1.0000
Near Poor/Low Income	High Income	-674.248250	291.591788	140	-2.31	0.0222	0.1333
Middle Income	High Income	-830.662639	311.350011	140	-2.67	0.0085	0.0512

Differences of TOTEXP18 Means for POVCAT18 Domains

95% Confidence Limits		Adjusted 95% Confidence Limits	
-123.5773	1398.79553	-392.7810	1667.99929
-20.3446	1608.39161	-308.3568	1896.40380
-847.3232	774.04497	-1134.0325	1060.75426
-495.6152	808.44398	-726.2142	1039.04297
-1250.7409	-97.75562	-1454.6252	106.12872
-1446.2183	-215.10694	-1663.9179	2.59257

Note: Due to the ODS EXCLUDE SUMMARY STATISTICS; statement, there is no output for ODS tables (SUMMARY and STATISTICS) here.

Ratio Analysis with PROC SURVEYMEANS

```
PROC SURVEYMEANS DATA=PUFMEPS.H209;
VAR TOTEXP18 OBVEXP18  OPTEXP18
      ERTEXP18 IPTEXP18 RXEXP18 HHAEXP18;
      STRATUM VARSTR;
      CLUSTER VARPSU;
  WEIGHT PERWT18F;
  RATIO OBVEXP18 OPTEXP18 ERTEXP18
        IPTEXP18 RXEXP18 HHAEXP18 / TOTEXP18;
run;
```

Code explanation: The RATIO statement requests ratio analysis for means or proportions of analysis variables.
https://support.sas.com/documentation/onlinedoc/stat/142/surveymeans.pdf

The above code block is part of the SAS program (referenced below), which includes the LIBNAME statement and global statements (e.g., OPTIONS and ODS GRAPHICS OFF statements) (not shown in this slide). See the SAS output in the next slide.

Source: ProcSurveyMeansRatio.SAS (https://github.com/pkmedu/AnalyzeMEPS).

Ratio Analysis with PROC SURVEYMEANS - Output

Statistics

Variable	N	Mean	Std Error of Mean	95% CL for Mean	
TOTEXP18	29415	6063.134458	128.011022	5810.04979	6316.21912
OBVEXP18	29415	1574.461264	38.187243	1498.96303	1649.95950
OPTEXP18	29415	547.080988	25.031525	497.59232	596.56966
ERTEXP18	29415	209.552169	8.572697	192.60349	226.50085
IPTEXP18	29415	1382.829611	72.633384	1239.22951	1526.42971
RXEXP18	29415	1389.881379	48.271278	1294.44647	1485.31629
HHAEXP18	29415	300.834870	23.249714	254.86894	346.80080

Ratio Analysis

Numerator	Denominator	N	Ratio	Std Error	95% CL for Ratio	
OBVEXP18	TOTEXP18	29415	0.259678	0.005224	0.24934881	0.27000674
OPTEXP18	TOTEXP18	29415	0.090231	0.003507	0.08329739	0.09716405
ERTEXP18	TOTEXP18	29415	0.034562	0.001353	0.03188706	0.03723632
IPTEXP18	TOTEXP18	29415	0.228072	0.009116	0.21004833	0.24609514
HHAEXP18	TOTEXP18	29415	0.049617	0.003949	0.04180876	0.05742535
RXEXP18	TOTEXP18	29415	0.229235	0.006553	0.21627889	0.24219070

The proportion of total health care spending (TOTEXP18) was 26.0 percent for medical providers' office-based care (OBEXP18), 22.9 percent for prescribed medicines (RXEXP18), and 22.8 percent for inpatient hospital care (IPTEXP18). The remaining distribution of total health care spending was 9.0 percent for outpatient hospital care (OPTEXP18), 5.0 percent for home health care (HHAEXP18), and 3.5 percent for emergency department care (ERTEXP18).

Questions?

Pradip.Muhuri@ahrq.hhs.gov