The MEPS Workshop Expounded

Scott Miner

Colorado State University – Global Campus

The MEPS Workshops Expounded

The Medical Expenditure Panel Survey (MEPS) is conducted to provide nationally representative estimates of health care expenditures, uses, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. MEPS is cosponsored by the Agency for Healthcare Research and Quality (AHRQ). MEPS consists of four component surveys: The Household Component (HC), the Medical Provider Component (MPC), the Insurance Component (IC), and the Nursing Home Component (NHC). The core of the survey lies within the HC, which forms the foundations for the MPC component and the IC component. The MEPS survey supports health services research, assesses health care policy implications, and yields comprehensive data on the distribution and level of healthcare use and expenditures (Moeller, Cohen, & Hock, 2002)

MEPS is the third survey in a series of surveys that started in 1977 with the introduction of the National Medical Care Expenditure Survey (NMCES) to assess the financing and use of medical care within America. Shortly after the introduction of NMCES, came the National Medical Expenditure Survey (NMES) in 1987. MEPS was officially introduced in 1996 and serves to provide a more current survey sample to help capture the ever-changing dynamic between U.S. citizens, medical care providers, and insurance providers (Moeller et al., 2002).

A few of the efficiencies incorporated into the MEPS survey include consolidating previous surveys, reducing costs, reducing burdens to respondents, enhancing longitudinal data collection, and enhancing analytical capabilities. MEPS respondents are a sub-sample of the respondents to the National Health Interview Survey (NHIS). MEPS enhances the data provided by NHIS by providing additional detailed data on health care expenditures and by linking this

data with data provided by medical providers, insurance providers, and employers (Moeller et al., 2002).

Household Component

The MEPS HC collects detailed healthcare expenditure data at both the person and household levels. The survey addresses topics such as demographic information, health condition, health status, access to care, satisfaction with care, charges and payments, insurance coverage status, income, and employment. The survey design is described as an overlapping panel design, where data is collected over a series of 5 interviews spanning two and a half years. Data is captured using computer-assisted personal interviewing or CAPI technology. Data is captured for two years for each household. Each year an additional panel is launched, adding new households into the mix. This structure allows the survey to capture the changing dynamics of the healthcare landscape. The survey design provides a nationally representative sample of the U.S. civilian noninstitutionalized population (Moeller et al., 2002).

Workshop

The MEPS workshops are conducted several times throughout the year. Workshops allow data analysts the opportunity to construct analytic scripts with the assistance of AHRQ staff. Additionally, workshops provide extended knowledge on MEPS data files and practical use information about the files. The workshops are designed for scientists and health services researchers who have a background or interest in data analytics ("Medical Expenditure Panel Survey Workshops and Events," n.d.).

In the following analyses the author provides the SAS script files along with detailed in-code commentary on each section of code. The script files are further modified in several

exercises to provide additional descriptive statistics and two-sample t-tests. After each script file, the author provides complementary output produced by each SAS program. Following the code and output, the author provides a report offering detailed analyses of each section of code. Lastly, additional screenshots are provided to demonstrate program execution and the use of version control via the GitHub repository.

Exercise 1.A Script

```
/*************************************************************************
DESCRIPTION:  THIS PROGRAM GENERATES THE FOLLOWING ESTIMATES ON NATIONAL
HEALTH CARE EXPENSES, 2016:
(1) OVERALL EXPENSES
(2) PERCENTAGE OF PERSONS WITH AN EXPENSE
(3) MEAN EXPENSE PER PERSON WITH AN EXPENSE
INPUT FILE:   C:\MEPS\SAS\DATA\H192.SAS7BDAT (2016 FULL-YEAR FILE)
 *************************************************************************/;

* Update log and print output information based on exercise;
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=1a;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
OPTIONS LINESIZE=160 PAGESIZE=79 NODATE FORMCHAR="|----|+|---+=|-/\<>*"
       PAGENO=1;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 "EXERCISE 1.A SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES, 2016";
/*************************************************************************
Create format variables
Age is divided into 3 categories: 'ALL AGES','0-64','65+'
A variable is created to represent whether a value is greater than zero
A FLAG is created with 3 possible outcomes: 'No or any expense',
'No expense', 'Any expense'
 *************************************************************************/;
PROC FORMAT;
       VALUE AGEF .='ALL AGES' 0-64='0-64' 65-HIGH='65+';
       VALUE AGECAT .='ALL AGES' 1='0-64' 2='65+';
       VALUE GTZERO 0='0' 0 <- HIGH='>0';
       VALUE FLAG .='No or any expense' 0='No expense' 1='Any expense';
RUN;
* Create summary statistics for dataset;
PROC MEANS DATA=CDATA.H192 n nmiss mean mode min max maxdec=2 range std fw=8;
       var TOTEXP16 AGE16X AGE42X AGE31X VARSTR VARPSU PERWT16F;
       title3 "Descriptive Statistics for age variables and expenditure
       variables in (2016 FULL-YEAR FILE) dataset";
RUN;
/*************************************************************************
READ IN DATA FROM 2016 CONSOLIDATED DATA FILE (HC-192)
The Age variable (AGE16X) represents the exact age, calculated from
date of birth and indicates age status as of 12/31/16.
Demographic variables that are round-specific are identified by names
including numbers “xy”, where x and y refer to round numbers of
```

```
Panel 20 and Panel 21 respectively.
For example, AGE31X represents the age data relevant to Round 3 of
Panel 20 or Round 1 of Panel 21. (It's Round #3 of Panel 20 because the
31 is marked with an X).
The variables VARSTR and VARPSU on this MEPS data file serve to identify
the sampling strata and primary sampling units required by the
variance estimation programs. In conjunction with the person-level
weight variable (PERWT16F) provided on this file, data for persons
with a positive person-level weight can be used to make estimates
for the civilian noninstitutionalized U.S. population for 2016.
Rename the TOTEXP16 Variable to Total
 ***************************************************************************/;
DATA PUF192;
       SET CDATA.H192 (KEEP=TOTEXP16 AGE16X AGE42X AGE31X VARSTR VARPSU PERWT16F);
       TOTAL=TOTEXP16;
       /* CREATE FLAG (1/0) VARIABLES FOR PERSONS WITH AN EXPENSE */
       X_ANYSVCE=0;
       IF TOTAL > 0 THEN
              X_ANYSVCE=1;
       /* CREATE A SUMMARY VARIABLE FROM END OF YEAR, 42, AND 31 VARIABLES*/
       IF AGE16X >=0 THEN
              AGE=AGE16X;
       ELSE IF AGE42X >=0 THEN
              AGE=AGE42X;
       ELSE IF AGE31X >=0 THEN
              AGE=AGE31X;
       /* CREATE 2 LEVEL AGE CATEGORIZATION VARIABLE (0-64,65+) */
       IF 0 LE AGE LE 64 THEN
              AGECAT=1;
       ELSE IF AGE  > 64 THEN
              AGECAT=2;
RUN;
/*****************************************************************************
Display the X_ANYSVCE FLAG Along side the TOTAL FLAG
Display the AGECAT FLAG Along side the AGE FLAG
Display Frequency counts for cross-tabs. Format as a LIST and treat missing
values as VALID non-missing levels.
 ***************************************************************************/;
PROC FREQ DATA=PUF192;
       TABLES X_ANYSVCE*TOTAL AGECAT*AGE /LIST MISSING;
       FORMAT TOTAL gtzero.
          AGE agef.;
       TITLE3 "Supporting crosstabs for the flag variables";
RUN;
* Suppress all output of the ODS;
ods graphics off;
ods exclude all;
/*****************************************************************************
Estimate the percentage of individuals with an expense & overall expenses
for the entire population
Estimate the proportion of individuals with an expense in the entire
population of the U.S.
Provide standard error estimations as well
Output the results of the statistics table to be later passed as input the
PROC PRINT statement for printing to the ODS Display
 ***************************************************************************/;
PROC SURVEYMEANS DATA=PUF192 MEAN NOBS SUMWGT STDERR SUM STD;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT16F;
       VAR X_ANYSVCE TOTAL;
       ods output Statistics=work.Overall_results;
       TITLE3 'PERCENTAGE OF PERSONS WITH AN EXPENSE & OVERALL EXPENSES';
RUN;
/* Exclude nothing from the output */
ods exclude none;
/*****************************************************************************
Print the results of the statistics table from the PROC SURVEYMEANS statement above
```

```
exclude the # of observations (NOOBS) and split the header on the * symbol

relabel variables for appropriate display output

format variables with commas for readability

Display the percentage of persons with an expense estimated for the entire
population, as well as the proportion
 ********************************************************************************/;
proc print data=work.Overall_results (firstobs=1 obs=1) noobs split='*';
       var N SumWgt mean StdErr Sum stddev;
       label SumWgt='Population*Size' mean='Proportion' StdErr='SE of Proportion'
              Sum='Persons*with Any*Expense '
              Stddev='SE of*Number*Persons*with*Any Expense';
       format N SumWgt Comma12. mean comma7.2 stderr 7.5 sum Stddev comma17.;
       TITLE3 'PERCENTAGE OF PERSONS WITH AN EXPENSE';
run;
/*******************************************************************************
Print additional results from the Statistics table from the PROC SURVEYMEANS
statement above
Begin printing with the 2nd row, exclude the # of observations, and split the
header on the '*' symbol
relabel variables for appropriate display output
format variables with commas for readability
Display the Overall Expenses estimates projected for the entire U.S. population
 ********************************************************************************/;
proc print data=work.Overall_results (firstobs=2) noobs split='*';
       var N SumWgt mean StdErr Sum stddev;
       label SumWgt='Population*Size' mean='Mean($)' StdErr='SE of Mean($)'
              Sum='Total*Expense ($)' Stddev='SE of*Total Expense($)';
       format N SumWgt Comma12. mean comma9.2 stderr 9.5 sum Stddev comma17.;
       TITLE3 'OVERALL EXPENSES';
run;
/* Exclude everything from the output */
ods exclude all;
/*******************************************************************************
Use the PROC SURVEYMEANS function to estimate the average expense per
person with an expense as well as estimate the average expense per person
with an expense as dependent upon age category (0-64, 65+)
print the output of the domain table to the domain_results dataset in the
temporary WORK library
The domain table will later be printed via the PROC PRINT procedure
 ********************************************************************************/;
PROC SURVEYMEANS DATA=PUF192 MEAN NOBS SUMWGT STDERR SUM STD;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT16F;
       VAR TOTAL;
       DOMAIN X_ANYSVCE('1') X_ANYSVCE('1')*AGECAT;
       FORMAT AGECAT agecat.;
       ods output domain=work.domain_results;
       TITLE3
              'MEAN EXPENSE PER PERSON WITH AN EXPENSE, FOR OVERALL, AGE 0-64, AND AGE 65+';
RUN;
/* Exclude nothing from the output */
ods exclude none;
/*******************************************************************************
Display the results of the PROC SURVEYMEANS procedure above
Display the AGECAT with user-defined formatting, the number of observations (N),
as well as Population Size and Mean($) Estimated Expenditure, etc.
 ********************************************************************************/;
proc print data=work.domain_results noobs split='*';
       var AGECAT N SumWgt mean StdErr Sum stddev;
       label AGECAT='Age Group' SumWgt='Population*Size' mean='Mean($)'
              StdErr='SE of Mean($)' Sum='Total*Expense ($)'
              Stddev='SE of*Total Expense($)';
       format AGECAT agecat. N SumWgt Comma12. mean comma9.1 stderr 9.4 sum Stddev
              comma17.;
run;
```

```
ods graphics on;
/*****************************************************************************
Perform a tTest on the variables AGECAT and TOTAL Expenditure

Display MEAN, STD DEV, STD ERR of the MEAN, MIN. and MAX. for each value of AGECAT
Display 95% CLI for MEAN and STD DEV in each AGECAT group.
Display 95% CLI for the difference in the group means using a pooled
estimate vs. Satterthwaite
Display the results of the t-test for both pooled and Satterthwaite versions
Display the results of the F-Test for deciding whether the variances
can be considered equal
Display Histogram (w/ normal curves, kernel density estimators, and
boxplots) and Q-Q Plots
 *****************************************************************************/;
PROC TTEST DATA=puf192;
       CLASS AGECAT;
       VAR TOTAL;
       WEIGHT PERWT16F;
       Format AGECAT agecat.;
       Where X_ANYSVCE = 1;
       Title3 'Independent Group T-Test Example (AGECAT By TOTAL Expenditure)';
Run;
/* Close All open ODS output destinations. */
ODS _ALL_ CLOSE;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
PROC PRINTTO;
RUN;
```

Exercise 1.A Output

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.A SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES, 2016
Descriptive Statistics for age variables and expenditure variables in (2016 FULL-YEAR FILE) dataset

The MEANS Procedure

Variable	Label	N	Miss	Mean	Mode	Minimum	Maximum	Range	N Std Dev
TOTEXP16	TOTAL HEALTH CARE EXP 16	34655	0	4325.05	0.00	0.00	580640.0	580640.0	13529.63
AGE16X	AGE AS OF 12/31/16 (EDITED/IMPUTED)	34655	0	36.12	7.00	-1.00	85.00	86.00	22.99
AGE42X	AGE - R4/2 (EDITED/IMPUTED)	34655	0	35.80	9.00	-1.00	85.00	86.00	23.12
AGE31X	AGE - R3/1 (EDITED/IMPUTED)	34655	0	35.20	-1.00	-1.00	85.00	86.00	23.32
VARSTR	VARIANCE ESTIMATION STRATUM - 2016	34655	0	1084.49	1144.00	1001.00	1165.00	164.00	48.55
VARPSU	VARIANCE ESTIMATION PSU - 2016	34655	0	1.63	2.00	1.00	3.00	2.00	0.62
PERWT16F	FINAL PERSON WEIGHT, 2016	34655	0	9324.53	0.00	0.00	99172.78	99172.78	8282.81

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.A SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES, 2016
Supporting crosstabs for the flag variables

The FREQ Procedure

X_ANYSVCE	TOTAL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	0	6933	20.01	6933	20.01
1	>0	27722	79.99	34655	100.00

AGECAT	AGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	0-64	29910	86.31	29910	86.31
2	65+	4745	13.69	34655	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.A SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES, 2016
PERCENTAGE OF PERSONS WITH AN EXPENSE

N	Population Size	Proportion	SE of Proportion	Persons with Any Expense	SE of Number Persons with Any Expense
33,259	323,141,687	0.85	0.00359	273,773,024	6,396,258

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.A SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES, 2016

OVERALL EXPENSES

N	Population Size	Mean($)	SE of Mean($)	Total Expense ($)	SE of Total Expense($)
33,259	323,141,687	5,005.64	116.95824	1,617,531,007,314	50,638,644,399

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.A SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES, 2016
MEAN EXPENSE PER PERSON WITH AN EXPENSE, FOR OVERALL, AGE 0-64, AND AGE 65+

Age Group	N	Population Size	Mean($)	SE of Mean($)	Total Expense ($)	SE of Total Expense($)
ALL AGES	26,942	273,773,024	5,908.3	133.8208	1,617,531,007,314	50,638,644,399
0-64	22,513	224,366,322	4,615.1	137.6071	1,035,478,590,271	38,245,335,117
65+	4,429	49,406,702	11,780.8	429.1720	582,052,417,044	26,901,634,384

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.A SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES, 2016
Independent Group T-Test Example (AGECAT By TOTAL Expenditure)

The TTEST Procedure

Variable: TOTAL

Weight: PERWT16F FINAL PERSON WEIGHT, 2016

AGECAT	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0-64		22513	4615.1	1443577	96.3743	1.0000	580640
65+		4429	11780.8	2092294	297.7	5.0000	236634
Diff (1-2)	Pooled		-7165.7	1568736	246.5		
Diff (1-2)	Satterthwaite		-7165.7		312.9		

AGECAT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0-64		4615.1	4426.2	4804.0	1443577	1430366	1457037
65+		11780.8	11197.3	12364.4	2092294	2049614	2136802
Diff (1-2)	Pooled	-7165.7	-7648.9	-6682.5	1568736	1555602	1582096
Diff (1-2)	Satterthwaite	-7165.7	-7779.1	-6552.3			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	26940	-29.07	<.0001
Satterthwaite	Unequal	5393.3	-22.90	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	4428	22512	2.10	<.0001

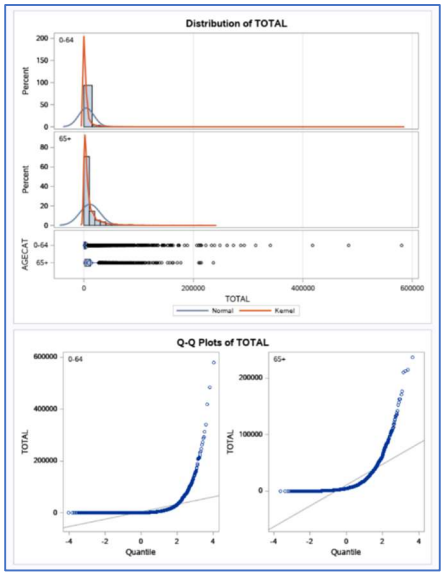

Figure 1. Histogram and Q-Q Plots for two sample t-test (expenditures by age group)

Exercise 1.B Script

```
/*************************************************************************
DESCRIPTION:  THIS PROGRAM GENERATES THE FOLLOWING ESTIMATES ON NATIONAL
HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015:
(1) PERCENTAGE DISTRIBUTION OF EXPENSES BY TYPE OF SERVICE
(2) PERCENTAGE OF PERSONS WITH AN EXPENSE, BY TYPE OF SERVICE
(3) MEAN EXPENSE PER PERSON WITH AN EXPENSE, BY TYPE OF SERVICE
DEFINED SERVICE CATEGORIES ARE:
HOSPITAL INPATIENT
AMBULATORY SERVICE: OFFICE-BASED & HOSPITAL OUTPATIENT VISITS
PRESCRIBED MEDICINES
DENTAL VISITS
EMERGENCY ROOM
HOME HEALTH CARE (AGENCY & NON-AGENCY) AND OTHER
(TOTAL EXPENDITURES - ABOVE EXPENDITURE CATEGORIES)
NOTE: EXPENSES INCLUDE BOTH FACILITY AND PHYSICIAN EXPENSES.
INPUT FILE:   C:\MEPS\SAS\DATA\H181.SAS7BDAT (2015 FULL-YEAR FILE)
 *************************************************************************/
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=1b;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
OPTIONS LINESIZE=160 PAGESIZE=79 NODATE FORMCHAR="|----|+|---+=|-/\<>*"
       PAGENO=1;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 "EXERCISE 1.B SAS: ESTIMATES ON NATIONAL
 HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015";
/*************************************************************************
Create format variables
Age is divided into 3 categories: 'ALL AGES','0-64','65+'
A variable is created to represent whether a value is greater than zero
A FLAG is created with 3 possible outcomes: 'No or any expense', 'No expense',
'Any expense'
 *************************************************************************/
PROC FORMAT;
       VALUE AGEF 0-64='0-64' 65-HIGH='65+';
       VALUE AGECAT .='All Ages' 1='0-64' 2='65+';
       VALUE GTZERO 0='$0' 0 <- HIGH='>$0';
       VALUE FLAG .='No or any expense' 0='No expense' 1='Any expense';
RUN;
/*************************************************************************
Create summary statistics for dataset
Display label information for the variables we are going to be analyzing,
along with more descriptive statistics for numerical variables
*************************************************************************/
       PROC MEANS DATA=CDATA.H181 n nmiss mean mode min max maxdec=2 range std fw=8;
var TOTEXP15 IPDEXP15 IPFEXP15 OBVEXP15 RXEXP15 OPDEXP15 OPFEXP15 DVTEXP15
       ERDEXP15 ERFEXP15 HHAEXP15 HHNEXP15 OTHEXP15 VISEXP15 AGE15X AGE42X AGE31X
       VARSTR VARPSU PERWT15f;
title3 "Descriptive Statistics for selected variables
 in the 2015 CONSOLIDATED DATA FILE (HC-181) ";
RUN;
/*************************************************************************
READ IN DATA FROM 2015 CONSOLIDATED DATA FILE (HC-181) (keep only specified fields)
Create TOTAL expenditure variables based on type of service
(I.E., Prescription Medicine, Dental, Hospital Inpatient, or
Home Health Care/Other)
Create DIFF variable to serve as a check (Calculate the DIFF of all Total
EXP Variables based on type - TOTAL)
Create ARRAY variables for FLAG variable creation
Do Loop iterates through Expenditure by Service Array
(I.E. Hospital_inpatient, ambulatory, etc.) and creates FLAG for
corresponding Expenditure by type FLAG
A Summary variable AGE is created from end-of-year and round-specific
```

```
AGE variables
 ***************************************************************************/
DATA PUF181;
       SET CDATA.H181 (KEEP=TOTEXP15 IPDEXP15 IPFEXP15 OBVEXP15 RXEXP15 OPDEXP15
              OPFEXP15 DVTEXP15 ERDEXP15 ERFEXP15 HHAEXP15 HHNEXP15 OTHEXP15 VISEXP15
              AGE15X AGE42X AGE31X VARSTR VARPSU PERWT15f);
       /* Define expenditure variables by type of service  */
       TOTAL=TOTEXP15;
       HOSPITAL_INPATIENT=IPDEXP15 + IPFEXP15;
       AMBULATORY=OBVEXP15 + OPDEXP15 + OPFEXP15 + ERDEXP15 + ERFEXP15;
       PRESCRIBED_MEDICINES=RXEXP15;
       DENTAL=DVTEXP15;
       HOME_HEALTH_OTHER=HHAEXP15 + HHNEXP15 + OTHEXP15 + VISEXP15;
       /*QC CHECK IF THE SUM OF EXPENDITURES BY TYPE OF SERVICE IS EQUAL
       TO TOTAL*/
       DIFF=TOTAL-HOSPITAL_INPATIENT - AMBULATORY   - PRESCRIBED_MEDICINES
              - DENTAL            - HOME_HEALTH_OTHER;
       /* CREATE FLAG (1/0) VARIABLES FOR PERSONS WITH AN EXPENSE,
       BY TYPE OF SERVICE*/
       ARRAY EXX  (6) TOTAL HOSPITAL_INPATIENT AMBULATORY PRESCRIBED_MEDICINES DENTAL
              HOME_HEALTH_OTHER;
       ARRAY ANYX (6) X_ANYSVCE X_HOSPITAL_INPATIENT X_AMBULATORY
              X_PRESCRIBED_MEDICINES X_DENTAL X_HOME_HEALTH_OTHER;
       /* Loop checks for any expense greater than 0 and assigns
       appropriate flag by type*/
       DO II=1 TO 6;
              ANYX(II)=0;
              IF EXX(II) > 0 THEN
                     ANYX(II)=1;
       END;
       DROP II;
       /* CREATE A SUMMARY VARIABLE FROM END OF YEAR, 42, AND 31 VARIABLES */
       IF AGE15X >=0 THEN
              AGE=AGE15X;
       ELSE IF AGE42X >=0 THEN
              AGE=AGE42X;
       ELSE IF AGE31X >=0 THEN
              AGE=AGE31X;
       /* Define AGECAT Variable based on (0-64)(65+) Age Group */
       IF 0 LE AGE LE 64 THEN
              AGECAT=1;
       ELSE IF AGE  > 64 THEN
              AGECAT=2;
       ;
RUN;
/***************************************************************************
Display the cross-tables for each FLAG Variable By Total Cost based on
Type of Service
QC Frequency Count Check
 ***************************************************************************/
PROC FREQ DATA=PUF181;
       TABLES X_ANYSVCE * TOTAL X_HOSPITAL_INPATIENT * HOSPITAL_INPATIENT
              X_AMBULATORY * AMBULATORY X_PRESCRIBED_MEDICINES * PRESCRIBED_MEDICINES
              X_DENTAL * DENTAL X_HOME_HEALTH_OTHER * HOME_HEALTH_OTHER AGECAT*AGE
              DIFF/LIST MISSING;
       FORMAT TOTAL HOSPITAL_INPATIENT AMBULATORY PRESCRIBED_MEDICINES DENTAL
              HOME_HEALTH_OTHER gtzero.
           AGE agef.
           X_ANYSVCE X_HOSPITAL_INPATIENT X_AMBULATORY
             X_PRESCRIBED_MEDICINES X_DENTAL X_HOME_HEALTH_OTHER flag.
           AGECAT agecat.;
       TITLE3 "Supporting crosstabs for the flag variables";
RUN;
/* Turn the graph output off */
ods graphics off;
/***************************************************************************
Display the ratios of Expense By Type of Service
Each type in the ratio option is divided by the total expenditure
 ***************************************************************************/
```

```
PROC SURVEYMEANS DATA=PUF181 sum;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT15f;
       VAR HOSPITAL_INPATIENT AMBULATORY PRESCRIBED_MEDICINES DENTAL
              HOME_HEALTH_OTHER TOTAL;
       RATIO HOSPITAL_INPATIENT AMBULATORY PRESCRIBED_MEDICINES DENTAL
              HOME_HEALTH_OTHER / TOTAL;
       TITLE3 'PERCENTAGE DISTRIBUTION OF EXPENSES BY
       TYPE OF SERVICE (STAT BRIEF #491 FIGURE 1)';
RUN;
/*****************************************************************************
Display percentage of persons with an expense by service type and
estimates for population at large
 *****************************************************************************/
PROC SURVEYMEANS DATA=PUF181 NOBS MEAN STDERR SUM;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT15f;
       VAR X_ANYSVCE X_HOSPITAL_INPATIENT X_AMBULATORY X_PRESCRIBED_MEDICINES
              X_DENTAL X_HOME_HEALTH_OTHER;
       TITLE3 'PERCENTAGE OF PERSONS WITH AN EXPENSE, BY TYPE OF SERVICE';
RUN;
/*****************************************************************************
Display the estimated average expenditure for the U.S. population
Display the estimated average expenditure for the U.S. population based on AGECAT
Display the estimated average expenditure for the U.S. population
for those w/ an EXPENSE, based on AGECAT
Output results of Statistics table and domain table to datasets in the
TEMPORARY WORK Library for later processing in PROC PRINT
 *****************************************************************************/
PROC SURVEYMEANS DATA=PUF181 MEAN NOBS SUMWGT STDERR SUM;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       VAR TOTAL;
       DOMAIN AGECAT AGECAT*X_ANYSVCE('1');
       WEIGHT PERWT15f;
       FORMAT AGECAT agecat.;
       ods output Statistics=work.Overall_results domain=work.domain_results;
       TITLE3 'MEAN TOTAL EXPENSE PER PERSON WITH AN
       EXPENSE, AGE 0-64, AND AGE 65+ (via ODS Output)';
RUN;
/* Create a new dataset based on the combined ODS OUTPUT above */
data combine;
       set work.Overall_results work.domain_results;
run;
/*****************************************************************************
Print the output of the combined results with appropriate formatting
and labels
Combine all the results above into one table that is more easily readable
with appropriate formatting and labels
 *****************************************************************************/
proc print data=combine noobs split='*';
       var AGECAT X_ANYSVCE N SumWgt mean StdErr Sum stddev;
       label AGECAT='Age Group' X_ANYSVCE='Expense*Category*(Flag)'
              SumWgt='Population*Size' mean='Mean($)' StdErr='SE of Mean($)'
              Sum='Total*Expense ($)' Stddev='SE of*Total Expense($)';
       format N SumWgt Comma12. mean comma7. stderr 7.3 sum Stddev comma17.
         X_ANYSVCE flag.;
       Title3 'MEAN TOTAL EXPENSE FOR ALL AGES,
       BY AGE GROUP, AND BY AGE GROUP WITH AN EXPENSE';
run;
/*****************************************************************************
The following perform a similar pattern analysis for each of the remaining
types of services.
The code in the workshop did not combine the results of the Statistics and
the domain tables, as it did in the above example with X_ANYSVCE.
Similar statistics to those above are generated for each type of service,
though the output is not combined into one table as it is above.
```

```
Display the estimated average expenditure for the U.S. population
Display the estimated average expenditure for the U.S. population per
person with an inpatient expense
Display the estimated average expenditure for the U.S. population
per person with an inpatient expense based on AGECAT
 ***************************************************************************/
PROC SURVEYMEANS DATA=PUF181 NOBS MEAN SUMWGT STDERR SUM;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT15f;
       VAR HOSPITAL_INPATIENT;
       DOMAIN X_HOSPITAL_INPATIENT('1') AGECAT*X_HOSPITAL_INPATIENT ('1');
       FORMAT AGECAT agecat.;
       TITLE3 'MEAN HOSPITAL INPATIENT EXPENSE PER PERSON WITH
       AN INPATIENT EXPENSE, AGE 0-64, AND AGE 65+';
RUN;
/***************************************************************************
Display the estimated average expenditure for the U.S. population.
Display the estimated average expenditure for the U.S. population
per person with an ambulance expense
Display the estimated average expenditure for the U.S. population per person
with an ambulance expense based on AGECAT
 ***************************************************************************/
PROC SURVEYMEANS DATA=PUF181 NOBS MEAN SUMWGT STDERR SUM;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT15f;
       VAR AMBULATORY;
       DOMAIN X_AMBULATORY('1') AGECAT*X_AMBULATORY('1');
       FORMAT AGECAT agecat.;
       TITLE3 'MEAN AMBULATORY EXPENSE PER PERSON WITH
       AN AMBULATORY EXPENSE, AGE 0-64, AND AGE 65+';
RUN;
/***************************************************************************
Display the estimated average expenditure for the U.S. population
Display the estimated average expenditure for the U.S. population per person
with a prescribed medicine expense
Display the estimated average expenditure for the U.S. population per person
with a prescribed medicine expense based on AGECAT
 ***************************************************************************/
PROC SURVEYMEANS DATA=PUF181 MEAN NOBS SUMWGT STDERR SUM
STRATUM VARSTR;
CLUSTER VARPSU;
WEIGHT PERWT15f;
VAR PRESCRIBED_MEDICINES;
DOMAIN X_PRESCRIBED_MEDICINES('1') AGECAT*X_PRESCRIBED_MEDICINES('1');
FORMAT AGECAT agecat.;
TITLE3 'MEAN PRESCRIPTION MEDICINE EXPENSE PER PERSON
WITH A PRESCRIPTION MEDICINE EXPENSE, AGE 0-64, AND AGE 65+';
RUN;
/***************************************************************************
Display the estimated average expenditure for the U.S. population
Display the estimated average expenditure for the U.S. population per person
with a dental expense
Display the estimated average expenditure for the U.S. population per person
with a dental expense based on AGECAT
 ***************************************************************************/
PROC SURVEYMEANS DATA=PUF181 MEAN NOBS SUMWGT STDERR SUM;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT15f;
       VAR DENTAL;
       DOMAIN X_DENTAL('1') AGECAT*X_DENTAL('1');
       FORMAT AGECAT agecat.;
       TITLE3
              'MEAN DENTAL EXPENSE PER PERSON WITH A DENATL
              EXPENSE, AGE 0-64, AND AGE 65+';
RUN;
/***************************************************************************
```

```
Display the estimated average expenditure for the U.S. population
Display the estimated average expenditure for the U.S. population per
person with a home health expense
Display the estimated average expenditure for the U.S. population per person
with a home health expense based on AGECAT
 ******************************************************************************/
PROC SURVEYMEANS DATA=PUF181 MEAN NOBS SUMWGT STDERR SUM;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT15f;
       VAR HOME_HEALTH_OTHER;
       DOMAIN X_HOME_HEALTH_OTHER('1') AGECAT*X_HOME_HEALTH_OTHER('1');
       FORMAT AGECAT agecat.;
       TITLE3 'MEAN OTHER EXPENSE (INCLUDING HOME HEALTH EXPENSE) PER PERSON WITH
       AN OTHER  EXPENSE, AGE 0-64, AND AGE 65+';
RUN;
ods graphics off;
ods select all;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 "EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH
CARE EXPENSES BY TYPE OF SERVICE, 2015";

/*****************************************************************************

Develop code for t-test.

Analyze AGE Category by Insurance Service Type

 *****************************************************************************/

PROC TTEST DATA=WORK.PUF181;
       CLASS AGECAT;
       VAR TOTAL HOSPITAL_INPATIENT AMBULATORY PRESCRIBED_MEDICINES DENTAL
              HOME_HEALTH_OTHER;
       TITLE3 "Additional T-Test for Type of Service By AGECAT";
       FORMAT AGECAT agecat.;
RUN;
PROC PRINTTO;
RUN;
```

Exercise 1.B Output

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
Supporting crosstabs for the flag variables

The FREQ Procedure

X_ANYSVCE	TOTAL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No expense	$0	6936	19.58	6936	19.58
Any expense	>$0	28491	80.42	35427	100.00

X_HOSPITAL_INPATIENT	HOSPITAL_INPATIENT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No expense	$0	33298	93.99	33298	93.99
Any expense	>$0	2129	6.01	35427	100.00

X_AMBULATORY	AMBULATORY	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No expense	$0	10515	29.68	10515	29.68
Any expense	>$0	24912	70.32	35427	100.00

X_PRESCRIBED_MEDICINES	PRESCRIBED_MEDICINES	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No expense	$0	15863	44.78	15863	44.78
Any expense	>$0	19564	55.22	35427	100.00

X_DENTAL	DENTAL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No expense	$0	22636	63.89	22636	63.89
Any expense	>$0	12791	36.11	35427	100.00

X_HOME_HEALTH_OTHER	HOME_HEALTH_OTHER	Frequency	Percent	Cumulative Frequency	Cumulative Percent
No expense	$0	29047	81.99	29047	81.99
Any expense	>$0	6380	18.01	35427	100.00

AGECAT	AGE	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0-64	0-64	30946	87.35	30946	87.35
65+	65+	4481	12.65	35427	100.00

DIFF	Frequency	Percent	Cumulative Frequency	Cumulative Percent
-3	1	0.00	1	0.00
-2	19	0.05	20	0.06
-1	766	2.16	786	2.22
0	33961	95.86	34747	98.08
1	665	1.88	35412	99.96
2	15	0.04	35427	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
PERCENTAGE DISTRIBUTION OF EXPENSES BY TYPE OF SERVICE (STAT BRIEF #491 FIGURE 1)

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	Sum	Std Error of Sum
HOSPITAL_INPATIENT	426203666990	24916792071
AMBULATORY	593594570455	22284974852
PRESCRIBED_MEDICINES	373012402939	17213224795
DENTAL	95918096998	3327595741
HOME_HEALTH_OTHER	111179693539	10281333639
TOTAL	1.5999074E12	52731878601

Ratio Analysis

Numerator	Denominator	Ratio	Std Error
HOSPITAL_INPATIENT	TOTAL	0.266393	0.011469
AMBULATORY	TOTAL	0.371018	0.009410
PRESCRIBED_MEDICINES	TOTAL	0.233146	0.006875
DENTAL	TOTAL	0.059952	0.001960
HOME_HEALTH_OTHER	TOTAL	0.069491	0.005958

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
PERCENTAGE OF PERSONS WITH AN EXPENSE, BY TYPE OF SERVICE

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	N	Mean	Std Error of Mean	Sum	Std Error of Sum
X_ANYSVCE	33893	0.853569	0.003762	274356895	6551954
X_HOSPITAL_INPATIENT	33893	0.063715	0.001927	20479364	703986
X_AMBULATORY	33893	0.760871	0.004463	244561763	5968022
X_PRESCRIBED_MEDICINES	33893	0.613064	0.004868	197052918	4821206
X_DENTAL	33893	0.417843	0.006402	134304354	3932674
X_HOME_HEALTH_OTHER	33893	0.212560	0.003794	68321692	2083517

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN TOTAL EXPENSE PER PERSON WITH AN EXPENSE, AGE 0-64, AND AGE 65+ (via ODS Output)

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
TOTAL	33893	321423251	4977.571930	134.414461	1.5999074E12	52731878601

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN TOTAL EXPENSE PER PERSON WITH AN EXPENSE, AGE 0-64, AND AGE 65+ (via ODS Output)

The SURVEYMEANS Procedure

Statistics for AGECAT Domains

AGECAT	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
0-64	TOTAL	29521	271829915	3876.446509	134.983704	1.0537341E12	42177372538
65+	TOTAL	4372	49593336	11013	336.115439	546173225160	23068225318

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN TOTAL EXPENSE PER PERSON WITH AN EXPENSE, AGE 0-64, AND AGE 65+ (via ODS Output)

The SURVEYMEANS Procedure

Statistics for AGECAT*X_ANYSVCE Domains

AGECAT	X_ANYSVCE	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
0-64	1	TOTAL	23448	226639193	4649.390571	159.062312	1.0537341E12	42177372538
65+	1	TOTAL	4184	47717702	11446	342.991434	546173225160	23068225318

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN TOTAL EXPENSE FOR ALL AGES, BY AGE GROUP, AND BY AGE GROUP WITH AN EXPENSE

Age Group	Category (Flag)	N	Population Size	Mean($)	SE of Mean($)	Total Expense ($)	Expense SE of Total Expense($)
All Ages	No or any expense	33,893	321,423,251	4,978	134.414	1,599,907,351,396	52,731,878,601
0-64	No or any expense	29,521	271,829,915	3,876	134.984	1,053,734,126,237	42,177,372,538
65+	No or any expense	4,372	49,593,336	11,013	336.115	546,173,225,160	23,068,225,318
0-64	Any expense	23,448	226,639,193	4,649	159.062	1,053,734,126,237	42,177,372,538
65+	Any expense	4,184	47,717,702	11,446	342.991	546,173,225,160	23,068,225,318

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN HOSPITAL INPATIENT EXPENSE PER PERSON WITH AN INPATIENT EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure
Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
HOSPITAL_INPATIENT	33893	321423251	1325.988913	79.053759	426203666990	24916792071

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN HOSPITAL INPATIENT EXPENSE PER PERSON WITH AN INPATIENT EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for X_HOSPITAL_INPATIENT Domains

X_HOSPITAL_ INPATIENT	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
1	HOSPITAL_INPATIENT	2045	20479364	20811	998.332312	426203666990	24916792071

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN HOSPITAL INPATIENT EXPENSE PER PERSON WITH AN INPATIENT EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for AGECAT*X_HOSPITAL_INPATIENT Domains

AGECAT	X_HOSPITAL_ INPATIENT	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
0-64	1	HOSPITAL_INPATIENT	1342	12233914	20482	1476.222296	250574288710	21638217854
65+	1	HOSPITAL_INPATIENT	703	8245450	21300	1279.161949	175629378281	12880405879

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN AMBULATORY EXPENSE PER PERSON WITHAN AMBULATORY EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
AMBULATORY	33893	321423251	1846.769233	51.144424	593594570455	22284974852

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN AMBULATORY EXPENSE PER PERSON WITHAN AMBULATORY EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for X_AMBULATORY Domains

X_AMBULATORY	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
1	AMBULATORY	24216	244561763	2427.176530	62.697077	593594570455	22284974852

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN AMBULATORY EXPENSE PER PERSON WITHAN AMBULATORY EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for AGECAT*X_AMBULATORY Domains

AGECAT	X_AMBULATORY	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
0-64	1	AMBULATORY	20191	198431386	2161.816655	67.265339	428972274932	18746444400
65+	1	AMBULATORY	4025	46130377	3568.631062	106.885564	164622295524	7569126221

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN DENTAL EXPENSE PER PERSON WITH A DENATL EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
DENTAL	33893	321423251	298.416797	7.614242	95918096998	3327595741

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN DENTAL EXPENSE PER PERSON WITH A DENATL EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for X_DENTAL Domains

X_DENTAL	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
1	DENTAL	12526	134304354	714.184565	17.100025	95918096998	3327595741

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN DENTAL EXPENSE PER PERSON WITH A DENATL EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for AGECAT*X_DENTAL Domains

AGECAT	X_DENTAL	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
0-64	1	DENTAL	10832	111592123	670.043229	18.377082	74771546322	2884075041
65+	1	DENTAL	1694	22712232	931.064418	45.620709	21146550675	1435722843

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN OTHER EXPENSE (INCLUDING HOME HEALTH EXPENSE) PER PERSON WITH AN OTHER EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
HOME_HEALTH_OTHER	33893	321423251	345.898105	31.126672	111179693539	10281333639

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN OTHER EXPENSE (INCLUDING HOME HEALTH EXPENSE) PER PERSON WITH AN OTHER EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for X_HOME_HEALTH_OTHER Domains

X_HOME_HEALTH_OTHER	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
1	HOME_HEALTH_OTHER	6237	68321692	1627.297129	144.112225	111179693539	10281333639

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
MEAN OTHER EXPENSE (INCLUDING HOME HEALTH EXPENSE) PER PERSON WITH AN OTHER EXPENSE, AGE 0-64, AND AGE 65+

The SURVEYMEANS Procedure

Statistics for AGECAT*X_HOME_HEALTH_OTHER Domains

AGECAT	HEALTH_OTHER	Variable	N	Sum of Weights	Mean	Std Error of Mean	Sum	X_HOME_ Std Error of Sum
0-64	1	HOME_HEALTH_OTHER	4699	50686433	1009.354848	163.964022	51160597066	8510015564
65+	1	HOME_HEALTH_OTHER	1538	17635259	3403.357819	275.860409	60019096474	5077547619

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 1.B SAS: ESTIMATES ON NATIONAL HEALTH CARE EXPENSES BY TYPE OF SERVICE, 2015
Additional T-Test for Type of Service By AGECAT

The TTEST Procedure

Variable: TOTAL

AGECAT	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0-64		30946	3227.5	12992.4	73.8564	0	700771
65+		4481	10687.6	18536.7	276.9	0	209436
Diff (1-2)	Pooled		-7460.1	13817.0	220.8		
Diff (1-2)	Satterthwaite		-7460.1		286.6		

AGECAT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0-64		3227.5	3082.8	3372.3	12992.4	12890.9	13095.6
65+		10687.6	10144.7	11230.5	18536.7	18160.8	18928.7
Diff (1-2)	Pooled	-7460.1	-7892.9	-7027.2	13817.0	13716.0	13919.5
Diff (1-2)	Satterthwaite	-7460.1	-8021.9	-6898.2			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	35425	-33.78	<.0001
Satterthwaite	Unequal	5136.3	-26.03	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	4480	30945	2.04	<.0001

Variable: HOSPITAL_INPATIENT

AGECAT	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0-64		30946	823.6	8839.4	50.2484	0	663916
65+		4481	3379.5	12346.1	184.4	0	184945
Diff (1-2)	Pooled		-2555.9	9355.8	149.5		
Diff (1-2)	Satterthwaite		-2555.9		191.2		

AGECAT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0-64		823.6	725.1	922.1	8839.4	8770.3	8909.6
65+		3379.5	3018.0	3741.1	12346.1	12095.7	12607.1
Diff (1-2)	Pooled	-2555.9	-2849.0	-2262.8	9355.8	9287.4	9425.2
Diff (1-2)	Satterthwaite	-2555.9	-2930.7	-2181.2			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	35425	-17.09	<.0001
Satterthwaite	Unequal	5165.6	-13.37	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	4480	30945	1.95	<.0001

Variable: AMBULATORY

AGECAT	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0-64		30946	1277.7	5553.5	31.5690	0	531259
65+		4481	3010.4	5782.8	86.3869	0	137642
Diff (1-2)	Pooled		-1732.6	5583.0	89.2366		
Diff (1-2)	Satterthwaite		-1732.6		91.9744		

AGECAT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0-64		1277.7	1215.9	1339.6	5553.5	5510.0	5597.6
65+		3010.4	2841.0	3179.7	5782.8	5665.5	5905.0
Diff (1-2)	Pooled	-1732.6	-1907.6	-1557.7	5583.0	5542.2	5624.4
Diff (1-2)	Satterthwaite	-1732.6	-1913.0	-1552.3			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	35425	-19.42	<.0001
Satterthwaite	Unequal	5741.6	-18.84	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	4480	30945	1.08	0.0003

Variable: PRESCRIBED_MEDICINES

AGECAT	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0-64		30946	738.9	4410.3	25.0707	0	414941
65+		4481	2531.8	5780.6	86.3551	0	129817
Diff (1-2)	Pooled		-1792.9	4606.2	73.6239		
Diff (1-2)	Satterthwaite		-1792.9		89.9208		

AGECAT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0-64		738.9	689.8	788.1	4410.3	4375.8	4445.3
65+		2531.8	2362.5	2701.1	5780.6	5663.4	5902.9
Diff (1-2)	Pooled	-1792.9	-1937.2	-1648.6	4606.2	4572.5	4640.4
Diff (1-2)	Satterthwaite	-1792.9	-1969.2	-1616.6			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	35425	-24.35	<.0001
Satterthwaite	Unequal	5261.6	-19.94	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	4480	30945	1.72	<.0001

Variable: DENTAL

AGECAT	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0-64		30946	223.2	817.3	4.6462	0	30000.0
65+		4481	365.6	1205.6	18.0099	0	25627.0
Diff (1-2)	Pooled		-142.4	876.0	14.0015		
Diff (1-2)	Satterthwaite		-142.4		18.5995		

AGECAT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0-64		223.2	214.1	232.3	817.3	810.9	823.8
65+		365.6	330.3	400.9	1205.6	1181.1	1231.1
Diff (1-2)	Pooled	-142.4	-169.9	-115.0	876.0	869.6	882.5
Diff (1-2)	Satterthwaite	-142.4	-178.9	-105.9			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	35425	-10.17	<.0001
Satterthwaite	Unequal	5092.9	-7.66	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	4480	30945	2.18	<.0001

Variable: HOME_HEALTH_OTHER

AGECAT	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
0-64		30946	164.1	2516.6	14.3059	0	201150
65+		4481	1400.3	6515.1	97.3268	0	172706
Diff (1-2)	Pooled		-1236.2	3301.6	52.7713		
Diff (1-2)	Satterthwaite		-1236.2		98.3726		

AGECAT	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0-64		164.1	136.1	192.1	2516.6	2496.9	2536.6
65+		1400.3	1209.5	1591.1	6515.1	6382.9	6652.8
Diff (1-2)	Pooled	-1236.2	-1339.6	-1132.7	3301.6	3277.4	3326.1
Diff (1-2)	Satterthwaite	-1236.2	-1429.0	-1043.3			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	35425	-23.43	<.0001
Satterthwaite	Unequal	4675.4	-12.57	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
Folded F	4480	30945	6.70	<.0001

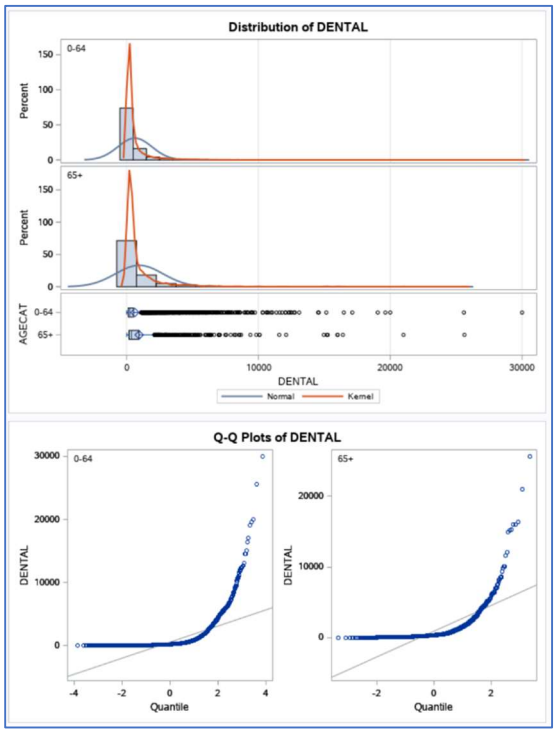

Figure 2. Histogram and Q-Q Plots for two sample t-test (dental expenditures by age group)

Exercise 2.A Script

```
/*************************************************************************\
PURPOSE:      THIS PROGRAM GENERATES SELECTED ESTIMATES FOR A 2015
VERSION OF THE MEPS STATISTICS BRIEF # 275: "Trends in Antipsychotics
Purchases and Expenses for the U.S. Civilian Noninstitutionalized
Population, 1997 and 2007"
(1) FIGURE 1: TOTAL EXPENSE FOR ANTIPSYCHOTICS
(2) FIGURE 2: TOTAL NUMBER OF PURCHASES OF ANTIPSYCHOTICS
(3) FIGURE 3: TOTAL NUMBER OF PERSONS PURCHASING ONE OR MORE
ANTIPSYCHOTICS
(4) FIGURE 4: AVERAGE TOTAL, OUT OF POCKET, AND THIRD PARTY
PAYER EXPENSE FOR ANTIPSYCHOTICS PER PERSON WITH AN ANTIPSYCHOTIC
MEDICINE PURCHASE
INPUT FILES:  (1) C:\MEPS\SAS\DATA\H1181.SAS7BDAT (2015 FULL-YEAR
CONSOLIDATED PUF) (2) C:\MEPS\SAS\DATA\H178A.SAS7BDAT
(2015 PRESCRIBED MEDICINES PUF)
 *************************************************************************/
```

```
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=2a;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
OPTIONS NODATE NONUMBER FORMCHAR="|----|+|---+=|-/\<>*";
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 "EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015";
/* Create a dataset with just a few variables for summary statistics */
DATA SUMMARYSET (KEEP=DUID PID DUPERSID RXRECIDX LINKIDX TC1 TC1S1 RXXP15X
              RXSF15X);
       SET CDATA.H178A;
RUN;
/* Select only the position table in the output */
ODS Select Position;
/*****************************************************************\
Display summary statistics:
Variables, Labels, Type of variable, size of variable:
Basic Summarized overview
 *****************************************************************/
PROC DATASETS;
       CONTENTS DATA=WORK.SUMMARYSET VARNUM;
       Title3 'Short list of labels for relevant variables';
       RUN;
       /* Select all output once again */
       ODS Select All;
       /*****************************************************************\

       Create summary statistics for the Sum of Payments and Amount Paid
       (Self or Family) Variables (2015 PRESCRIBED MEDICINES PUF)

        *****************************************************************/
PROC MEANS DATA=CDATA.H178A n nmiss mean mode min max maxdec=2 range std fw=8;
       var RXXP15X RXSF15X;
       title3 'Sum of Payments(2015), Amount Paid (Self or Family) (2015)';
RUN;
/*****************************************************************\
Create format variables
A variable is created to represent whether or
not a value is greater than zero
 *****************************************************************/
PROC FORMAT;
       VALUE GTZERO 0='0' 0 <- HIGH='>0';
RUN;
/*****************************************************************\
1) IDENTIFY Antipsychotic Drugs Using Therapeutic
Classification (Tc) Codes
Create DRUG dataset based on definition of antipsychotic drugs
 *****************************************************************/
DATA DRUG;
       SET CDATA.H178A;
       /*definition of ANTIPSYCHOTIC DRUGS*/
       IF TC1=242 AND TC1S1=251;
RUN;
/* Print a sample (30 observations) of our selected variables
by Person-ID (Including RX Expenditures) */
PROC PRINT DATA=DRUG (OBS=30);
       VAR RXRECIDX LINKIDX TC1 TC1S1 RXXP15X RXSF15X;
       BY DUPERSID;
       TITLE3 "A SAMPLE DUMP FOR PMED RECORDS WITH ANTIPSYCHOTIC DRUGS";
RUN;
/* 2) SUM DATA TO PERSON-LEVEL (Total out OOP Expenditures)
Output to a dataset named: PERDRUG */
PROC SUMMARY DATA=DRUG NWAY;
       CLASS DUPERSID;
       VAR RXXP15X RXSF15X;
```

```
	OUTPUT OUT=PERDRUG (DROP=_TYPE_) sum=TOT OOP;
RUN;
/* Print the first 30 values from PERDRUG
Will display the FREQ of RX per person along with TOTEXP and OOP */
PROC PRINT DATA=PERDRUG (OBS=30);
	TITLE3 "A SAMPLE DUMP FOR PERSON-LEVEL EXPENDITURES FOR ANTIPSYCHOTIC DRUGS";
RUN;
/* Create a new dataset where we rename _freq_ to 'N_PHRCHASE'
and create a calculated column called third payer,
based on total and out of pocket expense */
DATA PERDRUG2;
	SET PERDRUG;
	RENAME _FREQ_=N_PHRCHASE;
	THIRD_PAYER=TOT - OOP;
RUN;
/*****************************************************************\
3) MERGE THE PERSON-LEVEL EXPENDITURES TO THE ORIG DATASET
Get the variables from the original dataset that will allow us
to compute PROC SURVEYMEANS (STRATRA, CLUSTER, WEIGHT Variables)
Merge dataset with PERDRUG2 created above.
Create a bivariate FLAG (whether a person had an expenditure on
 antipsychotic medication based on the MERGE between datasets)
Additionally, we create appropriate labels for variables.
 *****************************************************************/
DATA FY;
	MERGE CDATA.H181 (IN=AA KEEP=DUPERSID VARSTR VARPSU PERWT15F) PERDRUG2  (IN=BB
		KEEP=DUPERSID N_PHRCHASE TOT OOP THIRD_PAYER);
	BY DUPERSID;
	IF AA AND BB THEN
		DO;
			SUB=1;
		END;
	ELSE IF NOT BB THEN
		DO;
			/*FOR PERSONS WITHOUT ANY PURCHASE OF ANTIPSYCHOTIC DRUGS*/
			SUB=2;
			N_PHRCHASE=0;
			THIRD_PAYER=0;
			TOT=0;
			OOP=0;
		END;
	IF AA;
	LABEL THIRD_PAYER='TOTAL-OOP' N_PHRCHASE='# OF PURCHASES PER PERSON'
		SUB='POPULATION FLAG FOR PERSONS WITH 1+ ANTIPSYCHOTIC DRUGS';
RUN;
/*****************************************************************\
Display Supporting cross-tabs for new variables
Population of N_PHRCHASE>0 with TOT>0 and OOP>0, etc. and FREQ Count
Format as LIST output, MISSING values are treated as valid nonmissing
levels
 *****************************************************************/
PROC FREQ DATA=FY;
	TABLES SUB * N_PHRCHASE * TOT * OOP * THIRD_PAYER / LIST MISSING;
	FORMAT N_PHRCHASE TOT OOP THIRD_PAYER gtzero.;
	TITLE3 "SUPPORTING CROSSTABS FOR NEW VARIABLES";
RUN;
/* Suppress Graph output */
ODS graphics off;
/*****************************************************************\
4) CALCULATE ESTIMATES ON EXPENDITURES AND USE
Display PERSON-LEVEL Estimates on EXPENDITURES and USE FOR
ANTIPSYCHOTIC DRUGS (2015)
Display Total results: Mean($) TOT Payment, N_PHRCHASE, OOP,
THIRD_PAYER
Display Results for Subpopulation (Person w/ ANTI-PSYCH EXP.:
Mean($) TOT Payment, N_PHRCHASE, OOP, THIRD_PAYER
 *****************************************************************/
PROC SURVEYMEANS DATA=FY NOBS SUMWGT SUM STD MEAN STDERR;
	STRATA VARSTR;
```

```
      CLUSTER VARPSU;
      WEIGHT PERWT15F;
      DOMAIN SUB('1');
      VAR TOT N_PHRCHASE OOP THIRD_PAYER;
      ODS OUTPUT DOMAIN=work.domain_results;
      TITLE3
            "PERSON-LEVEL ESTIMATES ON EXPENDITURES AND USE FOR ANTIPSYCHOTIC DRUGS, 2015";
RUN;
/* Close all output */
ODS _ALL_ CLOSE;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
PROC PRINTTO;
RUN;
```

Exercise 2.A Output

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015
Descriptive Summary Statistics

The DATASETS Procedure

Variables in Creation Order

#	Variable	Type	Len	Label
1	DUID	Num	6	DWELLING UNIT ID
2	PID	Num	4	PERSON NUMBER
3	DUPERSID	Char	8	PERSON ID (DUID + PID)
4	RXRECIDX	Char	15	UNIQUE RX/PRESCRIBED MEDICINE IDENTIFIER
5	LINKIDX	Char	12	ID FOR LINKAGE TO COND/OTH EVENT FILES
6	TC1	Num	4	MULTUM THERAPEUTIC CLASS #1
7	TC1S1	Num	4	MULTUM THERAPEUTIC SUB-CLASS #1 FOR TC1
8	RXSF15X	Num	8	AMOUNT PAID, SELF OR FAMILY (IMPUTED)
9	RXXP15X	Num	8	SUM OF PAYMENTS RXSF15X-RXOU15X(IMPUTED)

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015
Sum of Payments(2015), Amount Paid (Self or Family) (2015)

The MEANS Procedure

Variable	Label	N	N Miss	Mean	Mode	Minimum	Maximum	Range	Std Dev
RXXP15X	SUM OF PAYMENTS RXSF15X-RXOU15X(IMPUTED)	330453	0	103.53	4.00	0.00	32697.00	32697.00	492.89
RXSF15X	AMOUNT PAID, SELF OR FAMILY (IMPUTED)	330453	0	12.18	0.00	0.00	32697.00	32697.00	90.10

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015
A SAMPLE DUMP FOR PMED RECORDS WITH ANTIPSYCHOTIC DRUGS

PERSON ID (DUID + PID)=60001102

Obs	RXRECIDX	LINKIDX	TC1	TC1S1	RXXP15X	RXSF15X
1	600011020331001	600011020331	242	251	10	0
2	600011020601001	600011020601	242	251	10	0
3	600011020601002	600011020601	242	251	10	0
4	600011020601003	600011020601	242	251	10	0
5	600011020601004	600011020601	242	251	10	0
6	600011020601005	600011020601	242	251	10	0
7	600011020601006	600011020601	242	251	10	0
8	600011020601007	600011020601	242	251	10	0
9	600011020751001	600011020751	242	251	10	0

PERSON ID (DUID + PID)=60013101

Obs	RXRECIDX	LINKIDX	TC1	TC1S1	RXXP15X	RXSF15X
10	600131010501001	600131010501	242	251	6.70	1.2
11	600131010501002	600131010501	242	251	6.70	1.2
12	600131010501003	600131010501	242	251	6.70	1.2
13	600131010501004	600131010501	242	251	35.93	5.0
14	600131010501005	600131010501	242	251	35.93	5.0
15	600131010501006	600131010501	242	251	35.93	5.0
16	600131010581001	600131010581	242	251	40.93	5.0
17	600131010581002	600131010581	242	251	40.93	5.0
18	600131010581003	600131010581	242	251	40.93	5.0
19	600131010581004	600131010581	242	251	40.93	5.0
20	600131010581005	600131010581	242	251	40.93	5.0
21	600131010581006	600131010581	242	251	40.93	5.0

---------- PERSON ID (DUID + PID)=60047101 ----------

Obs	RXRECIDX	LINKIDX	TC1	TC1S1	RXXP15X	RXSF15X
22	600471010371001	600471010371	242	251	2.91	0
23	600471010721001	600471010721	242	251	2.91	0
24	600471010721002	600471010721	242	251	2.91	0
25	600471010721003	600471010721	242	251	2.91	0
26	600471010721004	600471010721	242	251	2.91	0
27	600471010721005	600471010721	242	251	2.91	0
28	600471010721006	600471010721	242	251	2.91	0
29	600471010721007	600471010721	242	251	2.91	0
30	600471010721008	600471010721	242	251	2.91	0

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015
A SAMPLE DUMP FOR PERSON-LEVEL EXPENDITURES FOR ANTIPSYCHOTIC DRUGS

Obs	DUPERSID	_FREQ_	TOT	OOP
1	60001102	9	90.00	0.00
2	60013101	12	373.47	48.60
3	60047101	14	357.76	0.00
4	60080101	11	60.98	4.15
5	60117105	4	1757.27	134.92
6	60141101	1	18.03	0.00
7	60187101	2	1163.08	4.80
8	60268101	1	37.00	37.00
9	60294101	1	497.99	0.00
10	60308101	4	11148.51	10.80
11	60325101	10	62.32	62.32
12	60430101	27	105.40	14.00
13	60449103	7	323.87	309.85
14	60494101	2	3927.82	0.00
15	60517102	11	134.87	5.50
16	60639101	10	4199.20	5.30
17	60698101	9	8678.20	215.00
18	60907101	6	238.38	21.90
19	60942101	15	16266.90	315.00
20	60968101	11	1154.89	29.15
21	61005101	9	25.05	0.00
22	61025101	12	464.28	0.00
23	61056101	1	22.00	4.01
24	61102102	2	1804.78	13.20
25	61146103	3	21.38	0.00
26	61191101	9	8071.12	0.00
27	61220101	11	1675.11	0.00
28	61297103	2	1272.70	40.00
29	61358101	9	1079.97	11.81
30	61371101	9	289.37	5.30

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015
SUPPORTING CROSSTABS FOR NEW VARIABLES

The FREQ Procedure

SUB	N_PHRCHASE	TOT	OOP	THIRD_PAYER	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	>0	0	0	0	1	0.00	1	0.00
1	>0	>0	0	>0	158	0.45	159	0.45
1	>0	>0	>0	0	45	0.13	204	0.58
1	>0	>0	>0	>0	255	0.72	459	1.30
2	0	0	0	0	34968	98.70	35427	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015
PERSON-LEVEL ESTIMATES ON EXPENDITURES AND USE FOR ANTIPSYCHOTIC DRUGS, 2015

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	Label	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
TOT	SUM OF PAYMENTS RXSF15X-RXOU15X (IMPUTED)	33893	321423251	31.860337	3.744940	10240653024	1205384106
N_PHRCHASE	# OF PURCHASES PER PERSON	33893	321423251	0.095843	0.006879	30806048	2267105
OOP	AMOUNT PAID, SELF OR	33893	321423251	1.714299	0.431995	551015438	138005747

```
                FAMILY (IMPUTED)
THIRD_PAYER     TOTAL-OOP                       33893       321423251       30.146038        3.643129      9689637586      1174371769
-------------------------------------------------------------------------------------------------------------------------------------

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2.A SAS: Antipsychotics Purchases and Expenses, 2015
PERSON-LEVEL ESTIMATES ON EXPENDITURES AND USE FOR ANTIPSYCHOTIC DRUGS, 2015

The SURVEYMEANS Procedure

Statistics for SUB Domains

                                                                     Sum of                          Std Error                         Std Error
SUB    Variable       Label                                N        Weights            Mean           of Mean             Sum            of Sum
------------------------------------------------------------------------------------------------------------------------------------------------
  1    TOT            SUM OF PAYMENTS                    440        4464275     2293.912000        226.463234     10240653024        1205384106
                      RXSF15X-RXOU15X
                      (IMPUTED)
       N_PHRCHASE     # OF PURCHASES PER                 440        4464275        6.900572          0.302250        30806048           2267105
                      PERSON
       OOP            AMOUNT PAID, SELF OR               440        4464275      123.427766         30.710067       551015438         138005747
                      FAMILY (IMPUTED)
       THIRD_PAYER    TOTAL-OOP                          440        4464275     2170.484234        221.282618      9689637586        1174371769
------------------------------------------------------------------------------------------------------------------------------------------------
```

Exercise 2.B Script

```
/*********************************************************************\
PURPOSE:     THIS PROGRAM GENERATES SELECTED ESTIMATES
FOR A 2016 VERSION OF THE Purchases and Expenses for Narcotic
analgesics or Narcotic analgesic combos
(1) FIGURE 1: TOTAL EXPENSE FOR Narcotic analgesics or Narcotic
analgesic combos
(2) FIGURE 2: TOTAL NUMBER OF PURCHASES OF Narcotic analgesics or
Narcotic analgesic combos
(3) FIGURE 3: TOTAL NUMBER OF PERSONS PURCHASING ONE OR MORE Narcotic
analgesics or Narcotic analgesic combos
(4) FIGURE 4: AVERAGE TOTAL, OUT OF POCKET, AND THIRD-PARTY PAYER
EXPENSE FOR Narcotic analgesics or Narcotic analgesic combos PER
PERSON WITH AN Narcotic analgesics or Narcotic analgesic combos
MEDICINE PURCHASE
INPUT FILES:  (1) C:\MEPS\SAS\DATA\H1192.SAS7BDAT
(2016 FULL-YEAR CONSOLIDATED PUF)(2) C:\MEPS\SAS\DATA\H188A.SAS7BDAT
(2016 PRESCRIBED MEDICINES PUF)
 *********************************************************************/
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=2b;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
OPTIONS LINESIZE=160 PAGESIZE=79 NODATE FORMCHAR="|----|+|---+=|-/\<>*"
       PAGENO=1;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 "EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic
       combos, 2016";
/* Create a dataset with just a few variables for summary statistics */
DATA SUMMARYSET (KEEP=DUID PID DUPERSID RXRECIDX LINKIDX TC1 TC1S1 RXXP16X
             RXSF16X);
       SET CDATA.H188A;
RUN;
/* Select only the position table in the output */
ODS Select Position;
/*********************************************************************\
Display summary statistics:
Variables, Labels, Type of variable, size of variable:
Basic Summarized overview
 *********************************************************************/
PROC DATASETS;
       CONTENTS DATA=WORK.SUMMARYSET VARNUM;
       Title3 'Short list of labels for relevant variables';
```

```
       RUN;
       /* Select all output once again */
       ODS Select All;
       /* Create summary statistics for the Sum of Payments
       and Amount Paid (Self or Family) Variables
       (2016 PRESCRIBED MEDICINES PUF) */
PROC MEANS DATA=CDATA.H188A n nmiss mean mode min max maxdec=2 range std fw=8;
       var RXXP16X RXSF16X;
       title3 'Sum of Payments, Amount Paid (Self or Family)';
RUN;
/* Create format variables
A variable is created to represent whether or
not a value is greater than zero */
PROC FORMAT;
       VALUE GTZERO 0='0' 0 <- HIGH='>0';
RUN;
/*****************************************************************\
1) IDENTIFY Narcotic analgesics or Narcotic analgesic combos
USING THERAPEUTIC CLASSIFICATION (TC) CODES
Create DRUG dataset based on definition of Narcotic analgesics
 *****************************************************************/
DATA DRUG;
       SET CDATA.H188A;
       /*definition of Narcotic analgesics or Narcotic analgesic combos*/
       IF TC1S1_1 IN (60, 191);
RUN;
/* Print a sample (30 observations) of our selected variables
by Person-ID (Including RX Expenditures) */
PROC PRINT DATA=DRUG (OBS=30);
       TITLE3 "A SAMPLE DUMP FOR PMED RECORDS WITH Narcotic
       analgesics or Narcotic analgesic combos";
       VAR RXRECIDX LINKIDX TC1S1_1 RXXP16X RXSF16X;
       BY DUPERSID;
RUN;
/* 2) SUM DATA TO PERSON-LEVEL (Total out OOP Expenditures)
Output to a dataset named: PERDRUG */
PROC SUMMARY DATA=DRUG NWAY;
       CLASS DUPERSID;
       VAR RXXP16X RXSF16X;
       OUTPUT OUT=PERDRUG (DROP=_TYPE_) sum=TOT OOP;
RUN;
/* Print the first 30 values from PERDRUG */
/* QC Check */
/* Will display the FREQ of RX per person along with TOTEXP and OOP */
PROC PRINT DATA=PERDRUG (OBS=30);
       TITLE3 "A SAMPLE DUMP FOR PERSON-LEVEL EXPENDITURES FOR Narcotic
       analgesics or Narcotic analgesic combos";
RUN;
/* Create a new dataset where we rename _freq_ to 'N_PHRCHASE'
and create a calculated column called third payer, based
on total and out of pocket expense */
DATA PERDRUG2;
       SET PERDRUG;
       RENAME _FREQ_=N_PHRCHASE;
       THIRD_PAYER=TOT - OOP;
RUN;
/*****************************************************************\
3) MERGE THE PERSON-LEVEL EXPENDITURES TO THE ORIG DATASET
Get the variables from the original dataset that will allow us
to compute PROC SURVEYMEANS (STRATRA, CLUSTER, WEIGHT Variables)
Merge dataset with PERDRUG2 created above.
Create a bivariate FLAG (whether a person had an expenditure on
Narcotic analgesics based on the MERGE between datasets)
Additionally, we create appropriate labels for variables.
 *****************************************************************/
DATA FY;
       MERGE CDATA.H192 (IN=AA KEEP=DUPERSID VARSTR VARPSU PERWT16F) PERDRUG2  (IN=BB
              KEEP=DUPERSID N_PHRCHASE TOT OOP THIRD_PAYER);
       BY DUPERSID;
```

```
      IF AA AND BB THEN
            DO;
                  SUB=1;
            END;
      ELSE IF NOT BB THEN
            DO;
                  /*FOR PERSONS WITHOUT ANY PURCHASE OF Narcotic
                  analgesics or Narcotic analgesic combos*/
                  SUB=2;
                  N_PHRCHASE=0;
                  THIRD_PAYER=0;
                  TOT=0;
                  OOP=0;
            END;
      IF AA;
      LABEL THIRD_PAYER='TOTAL-OOP' N_PHRCHASE='# OF PURCHASES PER PERSON'
      SUB='POPULATION FLAG FOR PERSONS WITH 1+ Narcotic analgesics or
      Narcotic analgesic combos';
RUN;
/*****************************************************************\
Display Supporting cross-tabs for new variables
Population of N_PHRCHASE>0 with TOT>0 and OOP>0, etc. and FREQ Count
Format as LIST output, MISSING values are treated as
valid nonmissing levels
 *****************************************************************/
PROC FREQ DATA=FY;
      TABLES SUB * N_PHRCHASE * TOT * OOP * THIRD_PAYER / LIST MISSING;
      FORMAT N_PHRCHASE TOT OOP THIRD_PAYER gtzero.;
      TITLE3 "SUPPORTING CROSSTABS FOR NEW VARIABLES";
RUN;
/*4) CALCULATE ESTIMATES ON USE AND EXPENDITURES*/
/* Suppress the printing of output */
ODS EXCLUDE ALL;
/*****************************************************************\
4) CALCULATE ESTIMATES ON EXPENDITURES AND USE
Calculate PERSON-LEVEL Estimates on EXPENDITURES and USE FOR
NARCOTIC ANALGESICS (2016)
Calculate Total results: Mean($) TOT Payment, N_PHRCHASE, OOP,
THIRD_PAYER
Calculate Results for Subpopulation (Person w/ NARCOTIC ANALGESICS
EXP.: Mean($) TOT Payment, N_PHRCHASE, OOP, THIRD_PAYER
Pipe output of Domain table to new temporary dataset for further
processing via PROC PRINT
 *****************************************************************/
PROC SURVEYMEANS DATA=FY NOBS SUMWGT SUM STD MEAN STDERR;
      STRATA VARSTR;
      CLUSTER VARPSU;
      WEIGHT PERWT16F;
      VAR TOT N_PHRCHASE OOP THIRD_PAYER;
      DOMAIN SUB('1');
      ODS OUTPUT DOMAIN=work.domain_results;
      TITLE3 "PERSON-LEVEL ESTIMATES ON EXPENDITURES AND USE FOR
      Narcotic analgesics or Narcotic analgesic combos, 2016";
RUN;
/* Unsuppress the printing of output */
ODS EXCLUDE NONE;
/* Display Results for Subpopulation (Person w/ NARCOTIC ANALGESICS
EXP) w/ projected estimates for entire U.S. Population */
proc print data=work.domain_results noobs split='*';
      var VARLABEL N SumWgt mean StdErr Sum stddev;
      label SumWgt='Population*Size' mean='Mean' StdErr='SE of Mean' Sum='Total'
            Stddev='SE of*Total';
      format N SumWgt Comma12. mean comma9.1 stderr 9.4 sum Stddev comma17.;
      TITLE3 "SUBSET THE ESTIMATES FOR PERSONS ONLY WITH 1+ Narcotic
      analgesics or Narcotic analgesic combos";
run;
/* Close all the output */
ODS _ALL_ CLOSE;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
```

```
PROC PRINTTO;
RUN;
```

Exercise 2.B Output

```
2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic  combos, 2016
Short list of labels for relevant variables

The DATASETS Procedure

Variables in Creation Order

#    Variable    Type    Len    Label

1    DUID        Num       6    DWELLING UNIT ID
2    PID         Num       4    PERSON NUMBER
3    DUPERSID    Char      8    PERSON ID (DUID + PID)
4    RXRECIDX    Char     15    UNIQUE RX/PRESCRIBED MEDICINE IDENTIFIER
5    LINKIDX     Char     12    ID FOR LINKAGE TO COND/OTH EVENT FILES
6    TC1         Num       4    MULTUM THERAPEUTIC CLASS #1
7    TC1S1       Num       4    MULTUM THERAPEUTIC SUB-CLASS #1 FOR TC1
8    RXSF16X     Num       8    AMOUNT PAID, SELF OR FAMILY (IMPUTED)
9    RXXP16X     Num       8    SUM OF PAYMENTS RXSF16X-RXOU16X(IMPUTED)
```

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic combos, 2016
Sum of Payments, Amount Paid (Self or Family)

The MEANS Procedure

Variable	Label	N	N Miss	Mean	Mode	Minimum	Maximum	Range	Std Dev
RXXP16X	SUM OF PAYMENTS RXSF16X-RXOU16X(IMPUTED)	319685	0	114.22	10.00	0.00	32132.50	32132.50	475.50
RXSF16X	AMOUNT PAID, SELF OR FAMILY (IMPUTED)	319685	0	12.14	0.00	0.00	6963.56	6963.56	67.82

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic combos, 2016
A SAMPLE DUMP FOR PMED RECORDS WITH Narcotic analgesics or Narcotic analgesic combos

PERSON ID (DUID + PID)=10002101

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
1	100021010211001	100021010211	191	88.68	0
2	100021010211002	100021010211	191	88.68	0
3	100021010211003	100021010211	191	88.68	0
4	100021010211004	100021010211	191	88.68	0
5	100021010211005	100021010211	191	88.68	0
6	100021010211006	100021010211	191	88.68	0

PERSON ID (DUID + PID)=10005101

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
7	100051010091001	100051010091	60	5	2.31

PERSON ID (DUID + PID)=10008107

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
8	100081070091001	100081070091	191	6.17	0

PERSON ID (DUID + PID)=10017101

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
9	100171010031001	100171010031	191	14.99	14.99

PERSON ID (DUID + PID)=10019102

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
10	100191020171001	100191020171	191	7.59	7.59

PERSON ID (DUID + PID)=10025102

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
11	100251020071001	100251020071	191	36.49	20
12	100251020121001	100251020121	191	36.49	20

PERSON ID (DUID + PID)=10028101

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
13	100281010031001	100281010031	191	5.97	5.97

PERSON ID (DUID + PID)=10029101

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
14	100291010121001	100291010121	191	7.22	0.28
15	100291010161001	100291010161	60	16.84	0.28
16	100291010161002	100291010161	60	4.81	0.20
17	100291010401001	100291010401	191	20.20	1.12
18	100291010401002	100291010401	191	20.20	1.12
19	100291010401003	100291010401	191	20.20	1.12
20	100291010401004	100291010401	191	20.20	1.12
21	100291010401005	100291010401	191	69.90	1.12
22	100291010401006	100291010401	191	69.90	1.12
23	100291010511001	100291010511	60	203.99	1.20
24	100291010761001	100291010761	191	15.12	1.12
25	100291010761002	100291010761	191	15.12	1.12
26	100291010761003	100291010761	191	15.12	1.12
27	100291010761004	100291010761	191	15.12	1.12

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic combos, 2016
A SAMPLE DUMP FOR PMED RECORDS WITH Narcotic analgesics or Narcotic analgesic combos

PERSON ID (DUID + PID)=10031102

Obs	RXRECIDX	LINKIDX	TC1S1_1	RXXP16X	RXSF16X
28	100311020221001	100311020221	60	1.34	0
29	100311020711001	100311020711	60	9.30	0
30	100311020711002	100311020711	60	7.08	0

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic combos, 2016
A SAMPLE DUMP FOR PERSON-LEVEL EXPENDITURES FOR Narcotic analgesics or Narcotic analgesic combos

Obs	DUPERSID	_FREQ_	TOT	OOP
1	10002101	6	532.08	0.00
2	10005101	1	5.00	2.31
3	10008107	1	6.17	0.00
4	10017101	1	14.99	14.99
5	10019102	1	7.59	7.59
6	10025102	2	72.98	40.00
7	10028101	1	5.97	5.97
8	10029101	14	513.94	13.16
9	10031102	3	17.72	0.00
10	10034102	6	132.15	132.15
11	10037101	12	459.88	6.00
12	10044102	1	12.98	2.35
13	10049101	8	393.57	63.53
14	10049102	10	234.70	0.00
15	10053102	1	4.38	0.00
16	10054101	1	5.00	0.18
17	10054102	2	95.00	3.72
18	10067101	12	37.39	0.00
19	10067102	11	33.75	0.00
20	10079101	1	262.00	117.90
21	10080101	4	196.28	92.78
22	10103101	5	24.20	16.83
23	10103102	1	21.97	21.97
24	10105101	1	20.99	0.37
25	10116101	7	355.85	72.50
26	10117101	8	17.15	6.00
27	10118102	1	262.00	29.44
28	10136102	1	10.68	0.52
29	10139101	1	65.99	1.00
30	10144101	1	9.40	0.90

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic combos, 2016
SUPPORTING CROSSTABS FOR NEW VARIABLES

The FREQ Procedure

SUB	N_PHRCHASE	TOT	OOP	THIRD_PAYER	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	>0	>0	0	>0	705	2.03	705	2.03
1	>0	>0	>0	0	957	2.76	1662	4.80
1	>0	>0	>0	>0	1592	4.59	3254	9.39
2	0	0	0	0	31401	90.61	34655	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 2b SAS: Narcotic analgesics or Narcotic analgesic combos, 2016
SUBSET THE ESTIMATES FOR PERSONS ONLY WITH 1+ Narcotic analgesics or Narcotic analgesic combos

Variable Label	N	Population Size	Mean	SE of Mean	Total	SE of Total
SUM OF PAYMENTS RXSF16X-RXOU16X(IMPUTED)	3,175	34,060,347	238.7	19.2872	8,131,321,798	680,339,817
# OF PURCHASES PER PERSON	3,175	34,060,347	3.6	0.0958	121,674,856	4,928,822
AMOUNT PAID, SELF OR FAMILY (IMPUTED)	3,175	34,060,347	41.8	3.7785	1,422,839,748	141,147,575
TOTAL-OOP	3,175	34,060,347	197.0	18.2409	6,708,482,050	629,279,347

Exercise 3.B Script

```
/*************************************************************************
DESCRIPTION:  THIS PROGRAM ILLUSTRATES HOW TO IDENTIFY PERSONS WITH A CONDITION AND
CALCULATE ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WITH THE CONDITION
THE CONDITION USED IN THIS EXERCISE IS DIABETES (CCS CODE=049 OR 050)
DEFINITION OF 61 CONDITIONS BASED ON THE CCS CODE
1  Infectious diseases                                          : CCS CODE = 1-9
2  Cancer                                                       : CCS CODE = 11-45
3  Non-malignant neoplasm                                       : CCS CODE = 46, 47
4  Thyroid disease                                              : CCS CODE = 48
5  Diabetes mellitus                                            : CCS CODE = 49,50
6  Other endocrine, nutritional & immune disorder               : CCS CODE = 51, 52, 54 - 58
7  Hyperlipidemia                                               : CCS CODE = 53
8  Anemia and other deficiencies                                : CCS CODE = 59
9  Hemorrhagic, coagulation, and disorders of White Blood cells : CCS CODE = 60-64
10 Mental disorders                                             : CCS CODE = 650-670
11 CNS infection                                                : CCS CODE = 76-78
12 Hereditary, degenerative and other nervous system disorders  : CCS CODE = 79-81
13 Paralysis                                                    : CCS CODE = 82
14 Headache                                                     : CCS CODE = 84
15 Epilepsy and convulsions                                     : CCS CODE = 83
16 Coma, brain damage                                           : CCS CODE = 85
17 Cataract                                                     : CCS CODE = 86
18 Glaucoma                                                     : CCS CODE = 88
19 Other eye disorders                                          : CCS CODE = 87, 89-91
20 Otitis media                                                 : CCS CODE = 92
21 Other CNS disorders                                          : CCS CODE = 93-95
22 Hypertension                                                 : CCS CODE = 98,99
23 Heart disease                                                : CCS CODE = 96, 97, 100-108
24 Cerebrovascular disease                                      : CCS CODE = 109-113
25 Other circulatory conditions arteries, veins, and lymphatics : CCS CODE = 114 -121
26 Pneumonia                                                    : CCS CODE = 122
27 Influenza                                                    : CCS CODE = 123
28 Tonsillitis                                                  : CCS CODE = 124
29 Acute Bronchitis and URI                                     : CCS CODE = 125 , 126
30 COPD, asthma                                                 : CCS CODE = 127-134
31 Intestinal infection                                         : CCS CODE = 135
32 Disorders of teeth and jaws                                  : CCS CODE = 136
33 Disorders of mouth and esophagus                             : CCS CODE = 137
34 Disorders of the upper GI                                    : CCS CODE = 138-141
35 Appendicitis                                                 : CCS CODE = 142
36 Hernias                                                      : CCS CODE = 143
37 Other stomach and intestinal disorders                       : CCS CODE = 144- 148
38 Other GI                                                     : CCS CODE = 153-155
39 Gallbladder, pancreatic, and liver disease                   : CCS CODE = 149-152
40 Kidney Disease                                               : CCS CODE = 156-158, 160, 161
41 Urinary tract infections                                     : CCS CODE = 159
42 Other urinary                                                : CCS CODE = 162,163
43 Male genital disorders                                       : CCS CODE = 164-166
44 Non-malignant breast disease                                 : CCS CODE = 167
45 Female genital disorders, and contraception                  : CCS CODE = 168-176
46 Complications of pregnancy and birth                         : CCS CODE = 177-195
47 Normal birth/live born                                       : CCS CODE = 196, 218
48 Skin disorders                                               : CCS CODE = 197-200
49 Osteoarthritis and other non-traumatic joint disorders       : CCS CODE = 201-204
50 Back problems                                                : CCS CODE = 205
51 Other bone and musculoskeletal  disease                      : CCS CODE = 206-209, 212
52 Systemic lupus and connective tissues disorders              : CCS CODE = 210-211
53 Congenital anomalies                                         : CCS CODE = 213-217
```

```
54 Perinatal Conditions                                    : CCS CODE = 219-224
55 Trauma-related disorders                                : CCS CODE = 225-236, 239, 240,
244
56 Complications of surgery or device                      : CCS CODE = 237, 238
57 Poisoning by medical and non-medical substances         : CCS CODE = 241 - 243
58 Residual Codes                                          : CCS CODE = 259
59 Other care and screening                                : CCS CODE = 10, 254-258
60 Symptoms                                                : CCS CODE = 245-252
61 Allergic reactions                                      : CCS CODE = 253
INPUT FILES:  1) C:\MEPS\SAS\DATA\H180.SAS7BDAT (2015 CONDITION PUF DATA)
2) C:\MEPS\SAS\DATA\H181.SAS7BDAT (2015 FY PUF DATA)
 ***************************************************************************/;
OPTIONS NODATE NONUMBER FORMCHAR="|----|+|---+=|-/\<>*";
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=3a;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
OPTIONS LINESIZE=160 PAGESIZE=79 NODATE FORMCHAR="|----|+|---+=|-/\<>*"
       PAGENO=1;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 'EXERCISE 3a SAS: CALCULATE ESTIMATES ON USE
 AND EXPENDITURES FOR PERSONS WITH A CONDITION (DIABETES)';
/* Create format variables to format output tables
Create format variables for SEX and YESNO */
PROC FORMAT;
       VALUE SEX
     .='TOTAL' 1='MALE' 2='FEMALE';
       VALUE YESNO
     .='TOTAL' 1='YES' 2='NO';
RUN;
/* Summary statistics table 1 */
DATA SUMMARYSET (KEEP=DUID PID CCCODEX);
       SET CDATA.H180;
RUN;
/* Summary statistics table 2 */
DATA SUMMARYSET2 (KEEP=TOTEXP15 TOTSLF15 OBTOTV15);
       SET CDATA.H181;
RUN;
/* Select only the position table in the output */
ODS Select Position;
/* Display summary statistics: File #1 */
PROC CONTENTS DATA=WORK.SUMMARYSET VARNUM;
       Title3 'Summary of H180 Variables';
RUN;
/* Select only the position table in the output */
ODS Select Position;
/* Display summary statistics: File #2 */
PROC CONTENTS DATA=WORK.SUMMARYSET2 VARNUM;
       Title3 'Summary of H181 Variables';
RUN;
/* Select all output once again */
ODS Select All;
/*****************************************************************\
Create a dataset named DIAB
1) PULL OUT CONDITIONS WITH DIABETES (CCS CODE='049', '050')
FROM 2015 CONDITION PUF - HC180
 *****************************************************************/
DATA DIAB;
       SET CDATA.H180;
       IF CCCODEX IN ('049', '050');
RUN;
/* Display summary counts for the values of the CCCODEX variable */
PROC FREQ DATA=DIAB;
       TABLES CCCODEX / LIST MISSING;
       TITLE3 "CHECK CCS CODES FOR DIABETIC CONDITIONS";
```

```
RUN;
/*****************************************************************\
2) IDENTIFY PERSONS WHO REPORTED DIABETES
SORT by DUPERSID
Remove any duplicates (NODUPKEY) as we consolidate records by DUPERSID
 *****************************************************************/
PROC SORT DATA=DIAB OUT=DIABPERS (KEEP=DUPERSID) NODUPKEY;
       BY DUPERSID;
RUN;
/*****************************************************************\
3) CREATE A FLAG FOR PERSONS WITH DIABETES IN THE 2015 FY DATA
Merge the original dataset with the dataset for the person
who reported diabetes and create a flag based on the result of the
merge and whether the DUPERSID exists in both tables.
Provide LABELS for out
 *****************************************************************/
DATA FY1;
       MERGE CDATA.H181 (IN=AA) DIABPERS   (IN=BB);
       BY DUPERSID;
       LABEL DIABPERS='PERSONS WHO REPORTED DIABETES';
       IF AA AND BB THEN
              DIABPERS=1;
       ELSE
              DIABPERS=2;
RUN;
/*****************************************************************\
Display FREQ count (UNWEIGHTED) for all persons and
FREQ by SEX (UNWEIGHTED)
Display as LIST format, MISSING values are treated as valid
nonmissing levels
Format variables as needed
 *****************************************************************/
PROC FREQ DATA=FY1;
       TABLES DIABPERS DIABPERS * SEX / LIST MISSING;
       FORMAT SEX sex.
         DIABPERS yesno.;
       TITLE3 "UNWEIGHTED # OF PERSONS WHO REPORTED DIABETES, 2015";
RUN;
/*****************************************************************\
Display FREQ count (WEIGHTED) for all persons and FREQ by
SEX (WEIGHTED)
Display as LIST format, MISSING values are treated as valid
nonmissing levels
Format variables as needed
 *****************************************************************/
PROC FREQ DATA=FY1;
       TABLES DIABPERS DIABPERS * SEX /LIST MISSING;
       WEIGHT PERWT15F;
       FORMAT SEX sex.
         DIABPERS yesno.;
       TITLE3 "WEIGHTED # OF PERSONS WHO REPORTED DIABETES, 2015";
RUN;
/* Turn off graphs */
ODS GRAPHICS OFF;
/* Suppress the listing output temporarily */
ODS LISTING CLOSE;
/*****************************************************************\
4) CALCULATE ESTIMATES ON EXPENDITURES AND USE
Calculate PERSON-LEVEL Estimates on USE and EXPENDITURES
entire population (2015)
Calculate PERSON-LEVEL Estimates on USE and EXPENDITURES
for persons with a condition (DIABETES) (2015)
Calculate PERSON-LEVEL Estimates on USE and EXPENDITURES for
persons with a condition (DIABETES) (2015) Based on SEX
Pass the output of the domain table to a new TEMP dataset for
further processing with PROC PRINT
 *****************************************************************/
PROC SURVEYMEANS DATA=FY1 NOBS SUMWGT SUM STD MEAN STDERR;
       STRATA VARSTR;
```

```
       CLUSTER VARPSU;
       WEIGHT PERWT15F;
       DOMAIN DIABPERS('1') SEX*DIABPERS('1');
       VAR TOTEXP15 TOTSLF15 OBTOTV15;
       ods output domain=work.domain_results;
RUN;
/* Turn the listing ODS output on */
ODS LISTING;
/*****************************************************************\
Print the PROC SURVEYMENTS output, dropping the column title
'domain label' and further specifying formatting options
such as blank lines, labels, & suppressing observation numbers
FORMAT the output with commas and number of characters to the
left and right of the decimal and user defined format variables
PRINT ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WHO REPORTED
DIABETES FOR 2015 BY SEX
 *****************************************************************/
PROC PRINT DATA=work.domain_results (DROP=DOMAINLABEL) NOOBS LABEL BLANKLINE=3;
       VAR SEX VARNAME N SUMWGT SUM STDDEV MEAN STDERR;
       FORMAT N comma6.0 SUMWGT SUM STDDEV comma17.0 MEAN STDERR comma9.2 DIABPERS
              yesno.
        SEX sex.;
       TITLE3
              "ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WHO
              REPORTED DIABETES, 2015";
RUN;
/* Close all the output */
ODS _ALL_ CLOSE;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
PROC PRINTTO;
RUN;
```

Exercise 3.A Output

```
2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 3a SAS: CALCULATE ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WITH A CONDITION (DIABETES)
Summary of H180 Variables

The CONTENTS Procedure

Variables in Creation Order

#    Variable    Type    Len    Label

1    DUID        Char      5    DWELLING UNIT ID
2    PID         Num       4    PERSON NUMBER
3    CCCODEX     Char      3    CLINICAL CLASSIFICATION CODE – EDITED

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 3a SAS: CALCULATE ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WITH A CONDITION (DIABETES)
Summary of H181 Variables

The CONTENTS Procedure

Variables in Creation Order

#    Variable    Type    Len    Label

1    TOTEXP15    Num       8    TOTAL HEALTH CARE EXP 15
2    TOTSLF15    Num       8    TOTAL AMT PAID BY SELF/FAMILY 15
3    OBTOTV15    Num       4    # OFFICE-BASED PROVIDER VISITS 15
```

```
                                                                          2018 AHRQ MEPS DATA USERS WORKSHOP
                                  EXERCISE 3a SAS: CALCULATE ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WITH A CONDITION (DIABETES)
                                                                              CHECK CCS CODES FOR DIABETIC CONDITIONS

                                                                                                     The FREQ Procedure

                                                                                          CLINICAL CLASSIFICATION CODE - EDITED

                                                                                                  Cumulative    Cumulative
                                                           CCCODEX    Frequency     Percent     Frequency       Percent
                                                           -------------------------------------------------------------
                                                            049           3228       97.02          3228         97.02
                                                            050             99        2.98          3327        100.00
```

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 3a SAS: CALCULATE ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WITH A CONDITION (DIABETES)
UNWEIGHTED # OF PERSONS WHO REPORTED DIABETES, 2015

The FREQ Procedure

PERSONS WHO REPORTED DIABETES

DIABPERS	Frequency	Percent	Cumulative Frequency	Cumulative Percent
YES	3065	8.65	3065	8.65
NO	32362	91.35	35427	100.00

DIABPERS	SEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
YES	MALE	1393	3.93	1393	3.93
YES	FEMALE	1672	4.72	3065	8.65
NO	MALE	15620	44.09	18685	52.74
NO	FEMALE	16742	47.26	35427	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 3a SAS: CALCULATE ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WITH A CONDITION (DIABETES)
WEIGHTED # OF PERSONS WHO REPORTED DIABETES, 2015

The FREQ Procedure

PERSONS WHO REPORTED DIABETES

DIABPERS	Frequency	Percent	Cumulative Frequency	Cumulative Percent
YES	27700589	8.62	27700589	8.62
NO	2.9372E8	91.38	3.2142E8	100.00

DIABPERS	SEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
YES	MALE	13531921	4.21	13531921	4.21
YES	FEMALE	14168669	4.41	27700589	8.62
NO	MALE	1.4364E8	44.69	1.7134E8	53.31
NO	FEMALE	1.5009E8	46.69	3.2142E8	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 3a SAS: CALCULATE ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WITH A CONDITION (DIABETES)
ESTIMATES ON USE AND EXPENDITURES FOR PERSONS WHO REPORTED DIABETES, 2015

SEX	Variable Name	N	SumWgt	Sum	Std Error of Sum	Mean	Std Error of Mean
TOTAL	TOTEXP15	2,982	27,700,589	393,028,078,843	24,786,133,857	14,188.44	803.76
TOTAL	TOTSLF15	2,982	27,700,589	33,333,948,175	1,487,670,122	1,203.37	47.36
TOTAL	OBTOTV15	2,982	27,700,589	327,151,092	14,514,086	11.81	0.37
MALE	TOTEXP15	1,359	13,531,921	195,119,396,038	14,540,065,365	14,419.19	937.74
MALE	TOTSLF15	1,359	13,531,921	16,486,769,548	1,108,692,452	1,218.36	68.42
MALE	OBTOTV15	1,359	13,531,921	148,875,724	9,300,471	11.00	0.54
FEMALE	TOTEXP15	1,623	14,168,669	197,908,682,805	14,806,750,887	13,968.05	975.73
FEMALE	TOTSLF15	1,623	14,168,669	16,847,178,628	965,315,394	1,189.04	63.78
FEMALE	OBTOTV15	1,623	14,168,669	178,275,368	9,488,276	12.58	0.49

Exercise 3.B Script

```
/*************************************************************************************
DESCRIPTION:  THIS PROGRAM ILLUSTRATES HOW TO CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH
A CONDITION
THE CONDITION USED IN THIS EXERCISE IS DIABETES (CCS CODE=049 OR 050)
THE DEFINITION OF 61 CONDITIONS BASED ON CCS CODE
1  Infectious diseases                                              : CCS CODE = 1-9
2  Cancer                                                           : CCS CODE = 11-45
3  Non-malignant neoplasm                                           : CCS CODE = 46, 47
4  Thyroid disease                                                  : CCS CODE = 48
5  Diabetes mellitus                                                : CCS CODE = 49,50
6  Other endocrine, nutritional & immune disorder                   : CCS CODE = 51, 52, 54 - 58
7  Hyperlipidemia                                                   : CCS CODE = 53
8  Anemia and other deficiencies                                    : CCS CODE = 59
9  Hemorrhagic, coagulation, and disorders of White Blood cells     : CCS CODE = 60-64
10 Mental disorders                                                 : CCS CODE = 650-670
```

```
11 CNS infection                                                  : CCS CODE = 76-78
12 Hereditary, degenerative and other nervous system disorders    : CCS CODE = 79-81
13 Paralysis                                                      : CCS CODE = 82
14 Headache                                                       : CCS CODE = 84
15 Epilepsy and convulsions                                       : CCS CODE = 83
16 Coma, brain damage                                             : CCS CODE = 85
17 Cataract                                                       : CCS CODE = 86
18 Glaucoma                                                       : CCS CODE = 88
19 Other eye disorders                                            : CCS CODE = 87, 89-91
20 Otitis media                                                   : CCS CODE = 92
21 Other CNS disorders                                            : CCS CODE = 93-95
22 Hypertension                                                   : CCS CODE = 98,99
23 Heart disease                                                  : CCS CODE = 96, 97, 100-108
24 Cerebrovascular disease                                        : CCS CODE = 109-113
25 Other circulatory conditions arteries, veins, and lymphatics   : CCS CODE = 114 -121
26 Pneumonia                                                      : CCS CODE = 122
27 Influenza                                                      : CCS CODE = 123
28 Tonsillitis                                                    : CCS CODE = 124
29 Acute Bronchitis and URI                                       : CCS CODE = 125 , 126
30 COPD, asthma                                                   : CCS CODE = 127-134
31 Intestinal infection                                           : CCS CODE = 135
32 Disorders of teeth and jaws                                    : CCS CODE = 136
33 Disorders of mouth and esophagus                               : CCS CODE = 137
34 Disorders of the upper GI                                      : CCS CODE = 138-141
35 Appendicitis                                                   : CCS CODE = 142
36 Hernias                                                        : CCS CODE = 143
37 Other stomach and intestinal disorders                         : CCS CODE = 144- 148
38 Other GI                                                       : CCS CODE = 153-155
39 Gallbladder, pancreatic, and liver disease                     : CCS CODE = 149-152
40 Kidney Disease                                                 : CCS CODE = 156-158, 160, 161
41 Urinary tract infections                                       : CCS CODE = 159
42 Other urinary                                                  : CCS CODE = 162,163
43 Male genital disorders                                         : CCS CODE = 164-166
44 Non-malignant breast disease                                   : CCS CODE = 167
45 Female genital disorders, and contraception                    : CCS CODE = 168-176
46 Complications of pregnancy and birth                           : CCS CODE = 177-195
47 Normal birth/live born                                         : CCS CODE = 196, 218
48 Skin disorders                                                 : CCS CODE = 197-200
49 Osteoarthritis and other non-traumatic joint disorders         : CCS CODE = 201-204
50 Back problems                                                  : CCS CODE = 205
51 Other bone and musculoskeletal  disease                        : CCS CODE = 206-209, 212
52 Systemic lupus and connective tissues disorders                : CCS CODE = 210-211
53 Congenital anomalies                                           : CCS CODE = 213-217
54 Perinatal Conditions                                           : CCS CODE = 219-224
55 Trauma-related disorders                                       : CCS CODE = 225-236, 239, 240,
244
56 Complications of surgery or device                             : CCS CODE = 237, 238
57 Poisoning by medical and non-medical substances                : CCS CODE = 241 - 243
58 Residual Codes                                                 : CCS CODE = 259
59 Other care and screening                                       : CCS CODE = 10, 254-258
60 Symptoms                                                       : CCS CODE = 245-252
61 Allergic reactions                                             : CCS CODE = 253
INPUT FILES:  1) C:\MEPS\SAS\DATA\H181.SAS7BDAT    (2015 FY PUF DATA)
2) C:\MEPS\SAS\DATA\H180.SAS7BDAT    (2015 CONDITION PUF DATA)
3) C:\MEPS\SAS\DATA\H178A.SAS7BDAT   (2015 PMED PUF DATA)
4) C:\MEPS\SAS\DATA\H178D.SAS7BDAT   (2015 INPATIENT VISITS PUF DATA)
5) C:\MEPS\SAS\DATA\H178E.SAS7BDAT   (2015 EROM VISITS PUF DATA)
6) C:\MEPS\SAS\DATA\H178F.SAS7BDAT   (2015 OUTPATIENT VISITS PUF DATA)
7) C:\MEPS\SAS\DATA\H178G.SASBDAT   (2015 OFFICE-BASED VISITS PUF DATA)
8) C:\MEPS\SAS\DATA\H178H.SAS7BDAT   (2015 HOME HEALTH PUF DATA)
9) C:\MEPS\SAS\DATA\H178IF1.SAS7BDAT  (2015 CONDITION-EVENT LINK PUF DATA)
 ************************************************************************/;
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=3b;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
OPTIONS LINESIZE=160 PAGESIZE=79 NODATE FORMCHAR="|----|+|---+=|-/\<>*"
      PAGENO=1;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
```

```
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 'AHRQ MEPS DATA USERS WORKSHOP -- 2018';
TITLE2 'EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED
WITH A CONDITION (DIABETES)';
/* Create format variables: greater than zero and greater than or equal to zero */
PROC FORMAT;
       VALUE GTZERO 0='0' 0 <- HIGH='>0';
       VALUE GEZERO 0 - HIGH='GE 0';
RUN;
/*******************************************************************\
1) PULL OUT CONDITIONS WITH DIABETES (CCS CODE='049', '050')
FROM 2015 CONDITION PUF - HC180
Create a dataset named DIAB
 *******************************************************************/
DATA DIAB;
       SET CDATA.H180;
       IF CCCODEX IN ('049', '050');
RUN;
/* Provide a frequency count of Diabetes condition codes */
PROC FREQ DATA=DIAB;
       TABLES CCCODEX / LIST MISSING;
TITLE3 "CHECK CCS CODES";
RUN;
/*******************************************************************\
2) GET EVENT ID FOR THE DIABETIC CONDITIONS FROM
CONDITION-EVENT LINK FILE
Merge Diabetes Dataset with the CONDITION-EVENT LINK PUF DATASET
CONDITION-EVENTLINK links an event (outpatient visit, inpatient visit,
ER visit, home health visit, etc.) to a condition (DIABETES)
Add the EVNTIDX field from the linking file, which corresponds
to a unique record in one of the MEPS event files, HC-178B through
HC-178H.
Call the dataset DIAB2
 *******************************************************************/
DATA DIAB2;
       MERGE DIAB          (IN=AA KEEP=DUPERSID CONDIDX CCCODEX) CDATA.H178IF1 (IN=BB
              KEEP=CONDIDX EVNTIDX);
       BY CONDIDX;
       IF AA AND BB;
RUN;
/* Review the output we just created by CONDIDX
(first 20 observations) */
PROC PRINT DATA=DIAB2 (OBS=20);
       BY CONDIDX;
TITLE3 "SAMPLE DUMP FOR CONDITION-EVENT LINK FILE";
RUN;
/*3) DELETE DUPLICATE CASES PER EVENT*/
/* CALL THE OUTPUT DIAB3 */
PROC SORT DATA=DIAB2 (KEEP=DUPERSID EVNTIDX) OUT=DIAB3 NODUPKEY;
       BY EVNTIDX;
RUN;
/* Review the output after deleting duplicated cases
first 30 observations) */
PROC PRINT DATA=DIAB3 (OBS=30);
TITLE3 "SAMPLE DUMP AFTER DUPLICATE CASES ARE DELETED";
RUN;
/*******************************************************************\
4) SUM UP PMED PURCHASE-LEVEL DATA TO EVENT-LEVEL
(EVNTIDX is not included on the 2015 Prescribed Medicines
event file, HC-178A;
Rather, on this file the variable for linking with EVNTIDX on the
CLNK file is LINKIDX.)
Include additional fields representing the event type
(outpatient visit, inpatient visit, ER visit, etc.) (RXSF15X--RXOUT15X)
Rename fields as discussed above
 *******************************************************************/
PROC SORT DATA=CDATA.H178A OUT=PMED (KEEP=LINKIDX RXXP15X RXSF15X--RXOU15X
```

```
             RENAME=(LINKIDX=EVNTIDX));
       BY LINKIDX;
RUN;
/* Summarize the data to the EVENTID Level; Call the output PMED2 */
PROC SUMMARY DATA=PMED NWAY;
       CLASS EVNTIDX;
       VAR RXXP15X RXSF15X--RXOU15X;
       OUTPUT OUT=PMED2 SUM=;
RUN;
/*******************************************************************\
5) ALIGN EXP VARIABLES IN DIFFERENT EVENTS WITH THE SAME NAMES
Prescribed Medications File (keep Event ID and source
of payment and total expenditure)
Rename variables to keep consistency among source of payment between
different event types
 *******************************************************************/
DATA PMED3 (KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP);
       SET PMED2;
       SF=RXSF15X;
       MR=RXMR15X;
       MD=RXMD15X;
       PV=RXPV15X;
       VA=RXVA15X;
       TR=RXTR15X;
       OF=RXOF15X;
       SL=RXSL15X;
       WC=RXWC15X;
       OR=RXOR15X;
       OU=RXOU15X;
       OT=RXOT15X;
       TOTEXP=RXXP15X;
       IF TOTEXP GE 0;
RUN;
/*******************************************************************\
Office-based Medical Provider Visits File (keep ID and source of
payment)
Rename variables to keep consistency among source of payment
between different event types
 *******************************************************************/
DATA OB (KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP);
       SET CDATA.H178G;
       SF=OBSF15X;
       MR=OBMR15X;
       MD=OBMD15X;
       PV=OBPV15X;
       VA=OBVA15X;
       TR=OBTR15X;
       OF=OBOF15X;
       SL=OBSL15X;
       WC=OBWC15X;
       OR=OBOR15X;
       OU=OBOU15X;
       OT=OBOT15X;
       TOTEXP=OBXP15X;
       IF TOTEXP GE 0;
RUN;

/*******************************************************************\
Emergency Room Event File (keep Event ID and source of payment and
total expenditure)
Rename variables to keep consistency among source of payment
between different event types
 *******************************************************************/
DATA EROM (KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP);
       SET CDATA.H178E;
       SF=ERFSF15X + ERDSF15X;
       MR=ERFMR15X + ERDMR15X;
       MD=ERFMD15X + ERDMD15X;
       PV=ERFPV15X + ERDPV15X;
```

```
       VA=ERFVA15X + ERDVA15X;
       TR=ERFTR15X + ERDTR15X;
       OF=ERFOF15X + ERDOF15X;
       SL=ERFSL15X + ERDSL15X;
       WC=ERFWC15X + ERDWC15X;
       OR=ERFOR15X + ERDOR15X;
       OU=ERFOU15X + ERDOU15X;
       OT=ERFOT15X + ERDOT15X;
       TOTEXP=ERXP15X;
       IF TOTEXP GE 0;
RUN;
/*****************************************************************\
Inpatient Event File (keep Event ID and source of payment
and total expenditure)
Rename variables to keep consistency among source of payment between
different event types
 *****************************************************************/
DATA IPAT (KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP);
       SET CDATA.H178D;
       SF=IPFSF15X + IPDSF15X;
       MR=IPFMR15X + IPDMR15X;
       MD=IPFMD15X + IPDMD15X;
       PV=IPFPV15X + IPDPV15X;
       VA=IPFVA15X + IPDVA15X;
       TR=IPFTR15X + IPDTR15X;
       OF=IPFOF15X + IPDOF15X;
       SL=IPFSL15X + IPDSL15X;
       WC=IPFWC15X + IPDWC15X;
       OR=IPFOR15X + IPDOR15X;
       OU=IPFOU15X + IPDOU15X;
       OT=IPFOT15X + IPDOT15X;
       TOTEXP=IPXP15X;
       IF TOTEXP GE 0;
RUN;
/*****************************************************************\
Home Health Visit Event File (keep Event ID and source of payment
and total expenditure)
Rename variables to keep consistency among source of payment between
different event types
 *****************************************************************/
DATA HVIS (KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP);
       SET CDATA.H178H;
       SF=HHSF15X;
       MR=HHMR15X;
       MD=HHMD15X;
       PV=HHPV15X;
       VA=HHVA15X;
       TR=HHTR15X;
       OF=HHOF15X;
       SL=HHSL15X;
       WC=HHWC15X;
       OR=HHOR15X;
       OU=HHOU15X;
       OT=HHOT15X;
       TOTEXP=HHXP15X;
       IF TOTEXP GE 0;
RUN;
/*****************************************************************\
Outpatient Event File (keep Event ID and source of
payment and total expenditure)
Rename variables to keep consistency among source of payment
between different event types
 *****************************************************************/
DATA OPAT (KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP);
       SET CDATA.H178F;
       SF=OPFSF15X + OPDSF15X;
       MR=OPFMR15X + OPDMR15X;
       MD=OPFMD15X + OPDMD15X;
       PV=OPFPV15X + OPDPV15X;
```

```
       VA=OPFVA15X + OPDVA15X;
       TR=OPFTR15X + OPDTR15X;
       OF=OPFOF15X + OPDOF15X;
       SL=OPFSL15X + OPDSL15X;
       WC=OPFWC15X + OPDWC15X;
       OR=OPFOR15X + OPDOR15X;
       OU=OPFOU15X + OPDOU15X;
       OT=OPFOT15X + OPDOT15X;
       TOTEXP=OPXP15X;
       IF TOTEXP GE 0;
RUN;
/*****************************************************************\
6)  COMBINE ALL EVENTS INTO ONE DATASET
Rename datasets with 2 letter references (i.e. IN=) and keep
corresponding variables created above
Create an EVNTYP variable to describe the type of event based
on the original file
Set the EVNTYP variable based on the name of the dataset that was
combined into ALLEVENT Dataset
Create labels for the display output
 *****************************************************************/
DATA ALLEVENT;
       SET OB   (IN=MV KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP)
              EROM (IN=ER KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP)
              IPAT (IN=ST KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP)
              HVIS (IN=HH KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP)
              OPAT (IN=OP KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP)
              PMED3 (IN=RX KEEP=EVNTIDX SF MR MD PV VA TR OF SL WC OR OU OT TOTEXP);
       BY EVNTIDX;
       LENGTH EVNTYP $4;
       LABEL EVNTYP='EVENT TYPE' TOTEXP='TOTAL EXPENDITURE FOR EVENT'
              SF="SOURCE OF PAYMENT: FAMILY" MR="SOURCE OF PAYMENT: MEDICARE"
              MD="SOURCE OF PAYMENT: MEDICAID" PV="SOURCE OF PAYMENT: PRIVATE INSURANCE"
              VA="SOURCE OF PAYMENT: VETERANS" TR="SOURCE OF PAYMENT: TRICARE"
              OF="SOURCE OF PAYMENT: OTHER FEDERAL"
              SL="SOURCE OF PAYMENT: STATE & LOCAL GOV"
              WC="SOURCE OF PAYMENT: WORKERS COMP" OR="SOURCE OF PAYMENT: OTHER PRIVATE"
              OU="SOURCE OF PAYMENT: OTHER PUBLIC" OT="SOURCE OF PAYMENT: OTHER INSURANCE";
       IF MV OR OP THEN
              EVNTYP='AMBU';
       ELSE IF ER THEN
              EVNTYP='EROM';
       ELSE IF ST THEN
              EVNTYP='IPAT';
       ELSE IF HH THEN
              EVNTYP='HVIS';
       ELSE IF RX THEN
              EVNTYP='PMED';
RUN;
/*****************************************************************\
Display frequency counts for the events combined into 1 file
LIST displays as list output, MISSING treats missing values as
valid nonmissing levels
FORMAT as appropriate
 *****************************************************************/
PROC FREQ DATA=ALLEVENT;
       TABLES EVNTYP TOTEXP SF MR MD PV VA TR OF SL WC OR OU OT /LIST MISSING;
       FORMAT TOTEXP SF MR MD PV VA TR OF SL WC OR OU OT gezero.;
TITLE3 "ALL EVENTS ARE COMBINED INTO ONE FILE";
RUN;
/* Print the first 20 observations (ALL EVENTS COMINTED INTO ONE FILE) */
PROC PRINT DATA=ALLEVENT (OBS=20);
RUN;
/*****************************************************************\
7) SUBSET EVENTS TO THOSE ONLY WITH DIABETES
MERGE DIAB3 dataset with ALLEVENT dataset
Create a dataset of all events for DIAB Condition called DIAB4
 *****************************************************************/
DATA DIAB4;
```

```
      MERGE DIAB3(IN=AA) ALLEVENT(IN=BB);
      BY EVNTIDX;
      IF AA AND BB;
RUN;
/*****************************************************************\
8) CALCULATE ESTIMATES ON EXPENDITURES AND USE, ALL TYPES OF SERVICE
Push output to ALL Dataset
Group by DUPERSID
 *****************************************************************/
PROC SUMMARY DATA=DIAB4 NWAY;
      CLASS DUPERSID;
      VAR TOTEXP SF MR MD PV VA TR OF SL WC OR OU OT;
      OUTPUT OUT=ALL SUM=;
RUN;
/*****************************************************************\
Merge the original dataset with the previous dataset to calculate
estimates on EXPENDITURES and USE for ALL TYPES OF SERVICE per person
Create a FLAG (Subpop.) for people who have at least 1 event
(Home health visit, ambulance, prescribed medication, emergency room)
associated with DIAB
Keep only those in Analysis with PERWT15F > 0
 *****************************************************************/
DATA FY1;
      MERGE CDATA.H181 (IN=AA KEEP=DUPERSID VARPSU VARSTR
            PERWT15F /*ADD MORE VARIABLES*/) ALL        (IN=BB KEEP=DUPERSID TOTEXP SF MR
            MD PV VA TR OF SL WC OR OU OT);
      BY DUPERSID;
      LABEL SUB='PERSONS WHO HAVE AT LEAST 1 EVENT ASSOCIATED WITH DIABETES';
      IF AA AND BB THEN
            SUB=1;
      ELSE IF AA AND NOT BB THEN
            DO;
                  /*PERSONS WITHOUT EVENTS WITH DIABETES*/
                  SUB=2;
                  TOTEXP=0;
                  SF=0;
                  MR=0;
                  MD=0;
                  PV=0;
                  VA=0;
                  TR=0;
                  OF=0;
                  SL=0;
                  WC=0;
                  OR=0;
                  OU=0;
                  OT=0;
            END;
      IF PERWT15F > 0;
RUN;
/* Turn off graphs and suppress output */
ODS GRAPHICS OFF;
/* Close the ODS LISTING output */
ODS LISTING CLOSE;
/*****************************************************************\
Calculate persons with condition (DIAB) who reported EVENT
(home health visit, prescribed medication, ambulance, etc. )
Calculate estimates for all sources of payments for sub-pop
(persons with DIAB reporting an Event)
 *****************************************************************/
PROC SURVEYMEANS DATA=FY1 NOBS SUMWGT SUM STD MEAN STDERR;
      STRATA VARSTR;
      CLUSTER VARPSU;
      WEIGHT PERWT15F;
      DOMAIN SUB('1');
      VAR TOTEXP SF MR MD PV VA TR OF SL WC OR OU OT;
      ODS OUTPUT DOMAIN=OUT1;
RUN;
/* Turn output on */
```

```
ODS LISTING;
/*****************************************************************\
Display ESTIMATES ON USE AND EXPENDITURES FOR ALL EVENTS
ASSOCIATED WITH DIABETES FOR 2015 BY SOURCE OF PAYMENTS
 *****************************************************************/
PROC PRINT DATA=OUT1 NOOBS LABEL;
       VAR VARNAME /*VARLABEL*/
       N SUMWGT SUM STDDEV MEAN STDERR;
       FORMAT N comma6.0 SUMWGT SUM STDDEV comma17.0 MEAN STDERR comma9.2;
TITLE3 "ESTIMATES ON USE AND EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH DIABETES, 2015";
RUN;
/*****************************************************************\
9) CALCULATE ESTIMATES ON EXPENDITURES AND USE BY TYPE OF SERVICE
Further consolidate data by summing up to PERSON-EVENT TYPE LEVEL
How many event types per person? 1 AMB, 13 PMED for person with DB
Pass output TOS (Type of Service) Temporary dataset
 *****************************************************************/
PROC SUMMARY DATA=DIAB4 NWAY;
       CLASS DUPERSID EVNTYP;
       VAR TOTEXP SF MR MD PV VA TR OF SL WC OR OU OT;
       OUTPUT OUT=TOS SUM=;
RUN;
/* Drop the _TYPE_ Field and rename _FREQ_ to N_VISITS
Call the dataset TOS2 (TYPE OF SERVICE 2) */
DATA TOS2;
       SET TOS (DROP=_TYPE_ RENAME=(_FREQ_=N_VISITS));
       LABEL N_VISITS='# OF VISITS PER PERSON FOR EACH TYPE OF SERVICE';
RUN;
/* Print the results to check output
QC - Sample Dump */
PROC PRINT DATA=TOS2 (OBS=20);
       BY DUPERSID;
TITLE3 "SAMPLE DUMP AFTER DATA IS SUMMED UP TO PERSON-EVENT TYPE-LEVEL";
RUN;
/*****************************************************************\
Create dataset and FLAG variable for Persons who have at least 1
event associated with diabetes.
SubPop either has event associated with DIAB (Condition) or does not.
Keep only those individuals for analysis with a PERWT15F > 0.
 *****************************************************************/
DATA FYTOS;
       MERGE CDATA.H181 (IN=AA KEEP=DUPERSID VARPSU VARSTR
              PERWT15F /*ADD MORE VARIABLES*/) TOS2       (IN=BB);
       BY DUPERSID;
       IF AA AND BB THEN
              SUB=1;
       ELSE IF AA AND NOT BB THEN
              DO;
                     /*PERSONS WITHOUT EVENTS WITH DIABETES*/
                     SUB=2;
                     EVNTYP='NA';
                     N_VISITS=0;
                     TOTEXP=0;
                     SF=0;
                     MR=0;
                     MD=0;
                     PV=0;
                     VA=0;
                     TR=0;
                     OF=0;
                     SL=0;
                     WC=0;
                     OR=0;
                     OU=0;
                     OT=0;
              END;
       LABEL SUB='PERSONS WHO HAVE AT LEAST 1 EVENT ASSOCIATED WITH DIABETES';
       IF PERWT15F > 0;
RUN;
```

```
/* Turn off graphics and turn off ODS LISTING output */
ODS GRAPHICS OFF;
ODS LISTING CLOSE;
/*******************************************************************\
Compute estimates by SOURCE PAYMENT for persons who reported an
EVENT with DIAB by EVENT TYPE.
Output the Domain table for further processing in PROC PRINT
statement.
 *******************************************************************/
PROC SURVEYMEANS DATA=FYTOS NOBS SUMWGT SUM STD MEAN STDERR;
       STRATA VARSTR;
       CLUSTER VARPSU;
       WEIGHT PERWT15F;
       DOMAIN SUB('1') * EVNTYP;
       VAR N_VISITS TOTEXP SF MR MD PV VA TR OF SL WC OR OU OT;
       ODS OUTPUT DOMAIN=OUT2;
RUN;
/* Turn ODS LISTING output on */
ODS LISTING;
/* Sort data by event type */
PROC SORT DATA=OUT2;
       BY EVNTYP;
RUN;
/*******************************************************************\
DISPLAY THE OUT2 DATASETS WITH LABELS
DISPLAY THE ESTIMATED EXPENTIRUES AND POPULATION TOTALS FOR INDIVIDUALS
REPORTING AND EVENT WITH DIABETES BY EVENT TYPE
 *******************************************************************/
PROC PRINT DATA=OUT2 NOOBS LABEL;
       BY EVNTYP;
       VAR VARNAME /*VARLABEL*/
TITLE3 "ESTIMATES ON USE AND EXPENDITURES FOR EVENTS ASSOCIATED WITH DIABETES, BY TYPE OF SERVICE,
2015";
       N SUMWGT SUM STDDEV MEAN STDERR;
       FORMAT N comma6.0 SUMWGT SUM STDDEV comma17.0 MEAN STDERR comma9.2;
RUN;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
PROC PRINTTO;
RUN;
```

Exercise 3.B Output

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
CHECK CCS CODES

The FREQ Procedure

CLINICAL CLASSIFICATION CODE - EDITED

CCCODEX	Frequency	Percent	Cumulative Frequency	Cumulative Percent
049	3228	97.02	3228	97.02
050	99	2.98	3327	100.00

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
SAMPLE DUMP FOR CONDITION-EVENT LINK FILE

CONDITION ID=600011010051

Obs	DUPERSID	CCCODEX	EVNTIDX
1	60001101	049	600011011361
2	60001101	049	600011011401
3	60001101	049	600011011431
4	60001101	049	600011011441
5	60001101	049	600011011701
6	60001101	049	600011011721
7	60001101	049	600011011741
8	60001101	049	600011011991
9	60001101	049	600011012001
10	60001101	049	600011012031
11	60001101	049	600011012071
12	60001101	049	600011012081

---------- CONDITION ID=600011010121 ----------

Obs	DUPERSID	CCCODEX	EVNTIDX
13	60001101	050	600011011451
14	60001101	050	600011011701
15	60001101	050	600011011771

---------- CONDITION ID=600011020031 ----------

Obs	DUPERSID	CCCODEX	EVNTIDX
16	60001102	049	600011020281
17	60001102	049	600011020311
18	60001102	049	600011020521
19	60001102	049	600011020581
20	60001102	049	600011020731

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
SAMPLE DUMP AFTER DUPLICATE CASES ARE DELETED

Obs	DUPERSID	EVNTIDX
1	60001101	600011011361
2	60001101	600011011401
3	60001101	600011011431
4	60001101	600011011441
5	60001101	600011011451
6	60001101	600011011701
7	60001101	600011011721
8	60001101	600011011741
9	60001101	600011011771
10	60001101	600011011991
11	60001101	600011012001
12	60001101	600011012031
13	60001101	600011012071
14	60001101	600011012081
15	60001102	600011020281
16	60001102	600011020311
17	60001102	600011020521
18	60001102	600011020581
19	60001102	600011020731
20	60004103	600041030011
21	60004103	600041030041
22	60004103	600041030071
23	60004103	600041030081
24	60004103	600041030091
25	60020101	600201010081
26	60035101	600351010461
27	60036101	600361010281
28	60036101	600361010291
29	60036101	600361010361
30	60036101	600361010461

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
ALL EVENTS ARE COMBINED INTO ONE FILE

The FREQ Procedure

EVENT TYPE

EVNTYP	Frequency	Percent	Cumulative Frequency	Cumulative Percent
AMBU	185183	54.12	185183	54.12
EROM	7163	2.09	192346	56.22
HVIS	5933	1.73	198279	57.95
IPAT	2921	0.85	201200	58.80
PMED	140953	41.20	342153	100.00

TOTAL EXPENDITURE FOR EVENT

TOTEXP	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: FAMILY

SF	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: MEDICARE

MR	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: MEDICAID

MD	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: PRIVATE INSURANCE

PV	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: VETERANS

VA	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: TRICARE

TR	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: OTHER FEDERAL

OF	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: STATE & LOCAL GOV

SL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: WORKERS COMP

WC	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: OTHER PRIVATE

OR	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: OTHER PUBLIC

OU	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

SOURCE OF PAYMENT: OTHER INSURANCE

OT	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GE 0	342153	100.00	342153	100.00

Obs	EVNTIDX	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT	TOTEXP	EVNTYP
1	600011011311	0.00	1287.28	3.12	0	0	0	0	0	0	0	0	0	1290.40	AMBU
2	600011011361	6.60	8.22	0.00	0	0	0	0	0	0	0	0	0	14.82	PMED
3	600011011381	0.00	48.48	0.00	0	0	0	0	0	0	0	0	0	48.48	PMED
4	600011011391	0.00	5.13	0.00	0	0	0	0	0	0	0	0	0	5.13	PMED
5	600011011401	0.00	75.88	0.00	0	0	0	0	0	0	0	0	0	75.88	PMED
6	600011011411	0.00	165.84	0.00	0	0	0	0	0	0	0	0	0	165.84	PMED
7	600011011421	0.00	73.96	0.00	0	0	0	0	0	0	0	0	0	73.96	PMED
8	600011011431	14.40	3013.44	0.00	0	0	0	0	0	0	0	0	0	3027.84	PMED
9	600011011441	3.60	215.13	0.00	0	0	0	0	0	0	0	0	0	218.73	PMED
10	600011011451	0.00	56.87	0.00	0	0	0	0	0	0	0	0	0	56.87	PMED
11	600011011471	0.00	13.32	0.00	0	0	0	0	0	0	0	0	0	13.32	PMED
12	600011011491	0.00	23.99	0.00	0	0	0	0	0	0	0	0	0	23.99	PMED
13	600011011501	19.35	0.00	0.00	0	0	0	0	0	0	0	0	0	19.35	PMED
14	600011011511	6.60	834.06	0.00	0	0	0	0	0	0	0	0	0	840.66	PMED
15	600011011521	0.00	151.96	0.00	0	0	0	0	0	0	0	0	0	151.96	PMED
16	600011011541	0.00	81.12	0.00	0	0	0	0	0	0	0	0	0	81.12	PMED
17	600011011561	4.79	0.00	0.00	0	0	0	0	0	0	0	0	0	4.79	PMED
18	600011011571	2.50	0.00	0.00	0	0	0	0	0	0	0	0	0	2.50	PMED
19	600011011581	1.79	0.00	0.00	0	0	0	0	0	0	0	0	0	1.79	PMED
20	600011011601	0.00	843.87	16.00	0	0	0	0	0	0	0	0	0	859.87	EROM

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
ESTIMATES ON USE AND EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH DIABETES, 2015

Variable Name	N	SumWgt	Sum	Std Error of Sum	Mean	Std Error of Mean
TOTEXP	2,831	26,356,142	89,665,504,958	5,165,060,112	3,402.07	167.10
SF	2,831	26,356,142	7,754,063,159	456,932,722	294.20	15.20
MR	2,831	26,356,142	35,492,281,908	3,182,445,303	1,346.64	113.88
MD	2,831	26,356,142	12,612,562,776	1,842,313,175	478.54	68.34
PV	2,831	26,356,142	25,605,505,962	2,243,003,700	971.52	80.05
VA	2,831	26,356,142	1,864,147,550	339,111,593	70.73	12.93
TR	2,831	26,356,142	1,600,311,164	606,348,575	60.72	22.95
OF	2,831	26,356,142	771,599,798	462,532,096	29.28	17.30
SL	2,831	26,356,142	1,217,886,863	307,410,213	46.21	11.63
WC	2,831	26,356,142	24,985,573	16,515,091	0.95	0.63
OR	2,831	26,356,142	1,745,745,854	398,150,040	66.24	15.07
OU	2,831	26,356,142	81,350,794	22,344,564	3.09	0.85
OT	2,831	26,356,142	895,063,556	402,135,500	33.96	15.27

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
SAMPLE DUMP AFTER DATA IS SUMMED UP TO PERSON-EVENT TYPE-LEVEL

PERSON ID (DUID + PID)=60001101

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
1	AMBU	1	69.46	0.0	0.00	69.46	0	0	0	0	0	0	0	0	0
2	PMED	13	11576.18	87.6	11488.58	0.00	0	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60001102

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
3	AMBU	1	134.29	0	0	134.29	0	0	0	0	0	0	0	0	0
4	PMED	4	105.82	0	0	105.82	0	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60004103

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
5	PMED	5	2120.73	401.54	0	0	0	0	0	0	0	0	1268.23	450.96	0

PERSON ID (DUID + PID)=60020101

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
6	PMED	1	32.21	0	0	0	0	0	0	32.21	0	0	0	0	0

PERSON ID (DUID + PID)=60035101

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
7	PMED	1	275.96	0	275.96	0	0	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60036101

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
8	AMBU	2	208.74	0.6	34.14	0	174.00	0	0	0	0	0	0	0	0
9	PMED	8	5896.21	88.0	0.00	0	5808.21	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60045101

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
10	AMBU	1	61.37	25.00	0	0	36.37	0	0	0	0	0	0	0	0
11	PMED	2	12.00	5.76	0	0	6.24	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60052101

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
12	PMED	1	10	10	0	0	0	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60060102

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
13	AMBU	4	344.24	79.00	0	0	265.24	0	0	0	0	0	0	0	0
14	PMED	5	6119.73	1021.87	0	0	5097.86	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60069101

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
15	AMBU	4	260	260	0	0	0	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60121103

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
16	AMBU	4	650.24	0.0	518.08	132.16	0	0	0	0	0	0	0	0	0
17	PMED	3	2211.86	13.2	2198.66	0.00	0	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60122202

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
18	AMBU	3	222.79	0.00	222.79	0	0	0	0	0	0	0	0	0	0
19	PMED	4	232.90	60.48	172.42	0	0	0	0	0	0	0	0	0	0

PERSON ID (DUID + PID)=60128102

Obs	EVNTYP	N_VISITS	TOTEXP	SF	MR	MD	PV	VA	TR	OF	SL	WC	OR	OU	OT
20	AMBU	5	352.35	330.35	0	0	22	0	0	0	0	0	0	0	0

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
SAMPLE DUMP AFTER DATA IS SUMMED UP TO PERSON-EVENT TYPE-LEVEL

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35931
Sum of Weights	340155378

Statistics

Variable	Label	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
N_VISITS	# OF VISITS PER PERSON FOR EACH TYPE OF SERVICE	35931	340155378	0.586018	0.019258	199337132	8119111
TOTEXP	TOTAL EXPENDITURE FOR EVENT	35931	340155378	263.601609	13.850919	89665504958	5165060112
SF	SOURCE OF PAYMENT: FAMILY	35931	340155378	22.795651	1.295110	7754063159	456932722
MR	SOURCE OF PAYMENT: MEDICARE	35931	340155378	104.341381	9.200484	35492281908	3182445303
MD	SOURCE OF PAYMENT: MEDICAID	35931	340155378	37.078828	5.361990	12612562776	1842313175
PV	SOURCE OF PAYMENT:	35931	340155378	75.275911	6.222783	25605505962	2243003700

	PRIVATE INSURANCE						
VA	SOURCE OF PAYMENT: VETERANS	35931	340155378	5.480282	0.992595	1864147550	339111593
TR	SOURCE OF PAYMENT: TRICARE	35931	340155378	4.704648	1.771984	1600311164	606348575
OF	SOURCE OF PAYMENT: OTHER FEDERAL	35931	340155378	2.268375	1.350110	771599798	462532096
SL	SOURCE OF PAYMENT: STATE & LOCAL GOV	35931	340155378	3.580384	0.889767	1217886863	307410213
WC	SOURCE OF PAYMENT: WORKERS COMP	35931	340155378	0.073453	0.048509	24985573	16515091
OR	SOURCE OF PAYMENT: OTHER PRIVATE	35931	340155378	5.132201	1.163528	1745745854	398150040
OU	SOURCE OF PAYMENT: OTHER PUBLIC	35931	340155378	0.239158	0.065459	81350794	22344564
OT	SOURCE OF PAYMENT: OTHER INSURANCE	35931	340155378	2.631337	1.178558	895063556	402135500

AHRQ MEPS DATA USERS WORKSHOP -- 2018
EXERCISE 3b SAS: CALCULATE EXPENDITURES FOR ALL EVENTS ASSOCIATED WITH A CONDITION (DIABETES)
SAMPLE DUMP AFTER DATA IS SUMMED UP TO PERSON-EVENT TYPE-LEVEL

The SURVEYMEANS Procedure

Statistics for SUB*EVNTYP Domains

SUB	EVNTYP	Variable	Label	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
1	AMBU	N_VISITS	# OF VISITS PER PERSON FOR EACH TYPE OF SERVICE	1989	18348311	4.022958	0.165995	73814485	4134339
		TOTEXP	TOTAL EXPENDITURE FOR EVENT	1989	18348311	863.900109	119.775000	15851107936	2255347515
		SF	SOURCE OF PAYMENT: FAMILY	1989	18348311	60.790401	6.286833	1115401180	123560206
		MR	SOURCE OF PAYMENT: MEDICARE	1989	18348311	384.807141	112.085578	7060561128	2073821092
		MD	SOURCE OF PAYMENT: MEDICAID	1989	18348311	72.758896	8.031214	1335002849	153336649
		PV	SOURCE OF PAYMENT: PRIVATE INSURANCE	1989	18348311	235.790652	31.729925	4326360237	585450057
		VA	SOURCE OF PAYMENT: VETERANS	1989	18348311	57.691280	12.707988	1058537553	233961540
		TR	SOURCE OF PAYMENT: TRICARE	1989	18348311	9.908037	3.708778	181795747	68957785
		OF	SOURCE OF PAYMENT: OTHER FEDERAL	1989	18348311	8.077305	7.729679	148204897	142747145
		SL	SOURCE OF PAYMENT: STATE & LOCAL GOV	1989	18348311	3.735975	1.710387	68548827	31668370
		WC	SOURCE OF PAYMENT: WORKERS COMP	1989	18348311	0.092094	0.043049	1689774	792665
		OR	SOURCE OF PAYMENT: OTHER PRIVATE	1989	18348311	14.388780	2.387818	264009820	44085135
		OU	SOURCE OF PAYMENT: OTHER PUBLIC	1989	18348311	1.616926	0.707219	29667856	12970333
		OT	SOURCE OF PAYMENT: OTHER INSURANCE	1989	18348311	14.242623	7.400688	261328069	134806956
	EROM	N_VISITS	# OF VISITS PER PERSON FOR EACH TYPE OF SERVICE	85	709974	1.649057	0.158416	1170787	202775
		TOTEXP	TOTAL EXPENDITURE FOR EVENT	85	709974	1868.182973	540.107805	1326360573	476694718
		SF	SOURCE OF PAYMENT: FAMILY	85	709974	178.314859	74.793378	126598841	59723824
		MR	SOURCE OF PAYMENT: MEDICARE	85	709974	144.902668	47.992392	102877068	33761390
		MD	SOURCE OF PAYMENT: MEDICAID	85	709974	136.766023	51.642811	97100264	38357365
		PV	SOURCE OF PAYMENT: PRIVATE INSURANCE	85	709974	775.319096	313.404511	550456083	265560507
		VA	SOURCE OF PAYMENT: VETERANS	85	709974	72.865728	38.969778	51732743	29859934
		TR	SOURCE OF PAYMENT: TRICARE	85	709974	2.024945	2.062710	1437657	1437657
		OF	SOURCE OF PAYMENT: OTHER FEDERAL	85	709974	0	0	0	0
		SL	SOURCE OF PAYMENT: STATE & LOCAL GOV	85	709974	13.536990	13.058981	9610906	9268755
		WC	SOURCE OF PAYMENT: WORKERS COMP	85	709974	0	0	0	0
		OR	SOURCE OF PAYMENT: OTHER PRIVATE	85	709974	33.987077	17.305791	24129927	12064166
		OU	SOURCE OF PAYMENT: OTHER PUBLIC	85	709974	0.120852	0.121207	85802	85802
		OT	SOURCE OF PAYMENT: OTHER INSURANCE	85	709974	510.344734	472.147227	362331283	342235939
	HVIS	N_VISITS	# OF VISITS PER PERSON FOR EACH TYPE OF SERVICE	132	1030192	4.512088	0.396755	4648315	709760
		TOTEXP	TOTAL EXPENDITURE FOR EVENT	132	1030192	6578.533923	1451.051496	6777150288	1752021246
		SF	SOURCE OF PAYMENT: FAMILY	132	1030192	222.928169	132.396205	229658724	140027353
		MR	SOURCE OF PAYMENT: MEDICARE	132	1030192	2278.319855	688.824333	2347105942	797117811
		MD	SOURCE OF PAYMENT: MEDICAID	132	1030192	3784.473628	1402.434586	3898732885	1526452748
		PV	SOURCE OF PAYMENT:	132	1030192	36.340050	25.029435	37437214	24734191

		PRIVATE INSURANCE						
	VA	SOURCE OF PAYMENT: VETERANS	132	1030192	67.700206	62.743270	69744183	65812647
	TR	SOURCE OF PAYMENT: TRICARE	132	1030192	2.020034	2.015963	2081022	2081022
	OF	SOURCE OF PAYMENT: OTHER FEDERAL	132	1030192	0	0	0	0
	SL	SOURCE OF PAYMENT: STATE & LOCAL GOV	132	1030192	109.807506	50.666543	113122769	50674760
	WC	SOURCE OF PAYMENT: WORKERS COMP	132	1030192	0	0	0	0
	OR	SOURCE OF PAYMENT: OTHER PRIVATE	132	1030192	74.043628	62.220894	76279122	64716974
	OU	SOURCE OF PAYMENT: OTHER PUBLIC	132	1030192	2.900847	2.916526	2988428	2988428
	OT	SOURCE OF PAYMENT: OTHER INSURANCE	132	1030192	0	0	0	0
IPAT	N_VISITS	# OF VISITS PER PERSON FOR EACH TYPE OF SERVICE	48	492910	1.210295	0.086765	596566	123694
	TOTEXP	TOTAL EXPENDITURE FOR EVENT	48	492910	12206	2188.971421	6016316285	1295502888
	SF	SOURCE OF PAYMENT: FAMILY	48	492910	367.913510	133.408291	181348305	76405367
	MR	SOURCE OF PAYMENT: MEDICARE	48	492910	6737.643917	1760.190882	3321053110	1002156852
	MD	SOURCE OF PAYMENT: MEDICAID	48	492910	2917.958659	1275.743791	1438291456	635473685
	PV	SOURCE OF PAYMENT: PRIVATE INSURANCE	48	492910	1531.657710	408.538254	754969640	197400289
	VA	SOURCE OF PAYMENT: VETERANS	48	492910	92.057005	93.877732	45375833	45375833
	TR	SOURCE OF PAYMENT: TRICARE	48	492910	132.693887	134.425206	65406164	65406164
	OF	SOURCE OF PAYMENT: OTHER FEDERAL	48	492910	0	0	0	0
	SL	SOURCE OF PAYMENT: STATE & LOCAL GOV	48	492910	105.505386	106.323092	52004676	52004676
	WC	SOURCE OF PAYMENT: WORKERS COMP	48	492910	0	0	0	0
	OR	SOURCE OF PAYMENT: OTHER PRIVATE	48	492910	158.051443	113.657355	77905162	56950886
	OU	SOURCE OF PAYMENT: OTHER PUBLIC	48	492910	1.955739	1.545919	964004	749730
	OT	SOURCE OF PAYMENT: OTHER INSURANCE	48	492910	160.268429	104.728423	78997936	52106011
PMED	N_VISITS	# OF VISITS PER PERSON FOR EACH TYPE OF SERVICE	2615	24506882	4.860144	0.096754	119106978	4652343
	TOTEXP	TOTAL EXPENDITURE FOR EVENT	2615	24506882	2435.828810	131.884179	59694569875	3778922778
	SF	SOURCE OF PAYMENT: FAMILY	2615	24506882	248.952765	13.655540	6101056109	374616236
	MR	SOURCE OF PAYMENT: MEDICARE	2615	24506882	924.666158	66.983736	22660684661	1799253460
	MD	SOURCE OF PAYMENT: MEDICAID	2615	24506882	238.440584	31.027436	5843435322	791965227
	PV	SOURCE OF PAYMENT: PRIVATE INSURANCE	2615	24506882	813.497310	76.895813	19936282789	1991368746
	VA	SOURCE OF PAYMENT: VETERANS	2615	24506882	26.064402	6.261617	638757239	152047675
	TR	SOURCE OF PAYMENT: TRICARE	2615	24506882	55.069860	23.585147	1349590574	579990340
	OF	SOURCE OF PAYMENT: OTHER FEDERAL	2615	24506882	25.437544	13.059963	623394902	325661598
	SL	SOURCE OF PAYMENT: STATE & LOCAL GOV	2615	24506882	39.768408	11.012108	974599685	269255653
	WC	SOURCE OF PAYMENT: WORKERS COMP	2615	24506882	0.950582	0.674214	23295800	16499508
	OR	SOURCE OF PAYMENT: OTHER PRIVATE	2615	24506882	53.185950	15.129561	1303421824	370340286
	OU	SOURCE OF PAYMENT: OTHER PUBLIC	2615	24506882	1.944136	0.493868	47644704	12259541
	OT	SOURCE OF PAYMENT: OTHER INSURANCE	2615	24506882	7.851112	3.371535	192406267	82823147

Exercise 4.A Script

```
/*************************************************************************************
DESCRIPTION:  THIS PROGRAM ILLUSTRATES HOW TO POOL MEPS DATA FILES FROM
DIFFERENT YEARS
THE EXAMPLE USED IS POPULATION AGE 26-30 WHO ARE UNINSURED BUT HAVE HIGH INCOME
DATA FROM 2015 AND 2016 ARE POOLED.
VARIABLES WITH YEAR-SPECIFIC NAMES MUST BE RENAMED BEFORE COMBINING FILES.
IN THIS PROGRAM THE INSURANCE COVERAGE VARIABLES 'INSCOV15' AND 'INSCOV16' ARE
 RENAMED TO 'INSCOV'.
SEE HC-036 (1996-2015 POOLED ESTIMATION FILE) FOR
INSTRUCTIONS ON POOOLING AND CONSIDERATIONS FOR VARIANCE
ESTIMATION FOR PRE-2002 DATA.
INPUT FILE:     (1) C:\MEPS\SAS\DATA\H192.SAS7BDAT (2016 FULL-YEAR FILE)
(2) C:\MEPS\SAS\DATA\H181.SAS7BDAT (2015 FULL-YEAR FILE)
 *************************************************************************/;
```

```
OPTIONS NODATE NONUMBER PAGESIZE=200 FORMCHAR="|----|+|---+=|-/\<>*";
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=4a;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2
       'EXERCISE6.SAS: POOL MEPS DATA FILES FROM DIFFERENT YEARS (2015 and 2016)';
/* Create format variables */
PROC FORMAT;
       VALUE POVCAT 1='1 POOR/NEGATIVE' 2='2 NEAR POOR' 3='3 LOW INCOME'
              4='4 MIDDLE INCOME' 5='5 HIGH INCOME';
       VALUE INSF 1='1 ANY PRIVATE' 2='2 PUBLIC ONLY' 3='3 UNINSURED';
       VALUE AGE 26-30='26-30' 0-25='0-25' 31-HIGH='31+';
run;
/* Summary statistics table 2 */
DATA SUMMARYSET2 (KEEP=DUPERSID INSCOV15 PERWT15F VARSTR VARPSU POVCAT15
             AGELAST TOTSLF15);
       SET CDATA.H181;
RUN;
/* Select only the position table in the output */
ODS Select Position;
/* Display summary statistics: File #2 */
PROC CONTENTS DATA=WORK.SUMMARYSET2 VARNUM;
       Title3 'Summary of H181 Variables';
RUN;
/* Select all output once again */
ODS Select All;
/*****************************************************************************
FREQUENCY OF 2015
Create a dataset called YR1
Keep PERSON ID, AGE, INC AS % OF POVERTY LINE, and INSURANCE COVERAGE INDICATOR
 *****************************************************************************/;
DATA YR1;
       SET CDATA.H181 (KEEP=DUPERSID INSCOV15 PERWT15F VARSTR VARPSU POVCAT15 AGELAST
             TOTSLF15);
       IF PERWT15F>0;
RUN;
/*****************************************************************************
PROC FREQ show INC AS % OF POVERTY LINE by INSURANCE COVERAGE INDICATOR
(26 <= AGE <= 30)
Display UNWEIGHTED output
Display LIST format MISSING values are treated as valid nonmissing levels
Format variables
Display Poverty Category By Insurance Coverage Indicator with FREQ Count
 *****************************************************************************/;
PROC FREQ DATA=YR1 (WHERE=(26 LE AGELAST LE 30));
       TABLES POVCAT15*INSCOV15/ LIST MISSING;
       FORMAT INSCOV15 INSF.  POVCAT15 POVCAT.;
       TITLE3 'UNWEIGHTED FREQUENCY FOR 2015 FY PERSONS WITH AGE 26-30';
RUN;
/*****************************************************************************
FREQUENCY OF 2016
Create a dataset called YR2
Keep PERSON ID, AGE, INC AS % OF POVERTY LINE, and INSURANCE COVERAGE INDICATOR
 *****************************************************************************/;
DATA YR2;
       SET CDATA.H192 (KEEP=DUPERSID INSCOV16 PERWT16F VARSTR VARPSU POVCAT16 AGELAST
             TOTSLF16);
       IF PERWT16F>0;
run;
/*****************************************************************************
PROC FREQ show INC AS % OF POVERTY LINE by INSURANCE COVERAGE INDICATOR
(26 <= AGE <= 30)
Display UNWEIGHTED output
```

```
Display LIST format MISSING values are treated as valid nonmissing levels
Format variables
Display Poverty Category By Insurance Coverage Indicator with FREQ Count
 ********************************************************************************/;
PROC FREQ DATA=YR2 (WHERE=(26 LE AGELAST LE 30));
       TABLES POVCAT16*INSCOV16/ LIST MISSING;
       FORMAT INSCOV16 INSF.  POVCAT16 POVCAT.;
       TITLE3 'UNWEIGHTED FREQUENCY FOR 2016 FY PERSONS WITH AGE 26-30';
RUN;
/* RENAME YEAR SPECIFIC VARIABLES PRIOR TO COMBINING FILES */
DATA YR1X;
       SET YR1 (RENAME=(INSCOV15=INSCOV PERWT15F=PERWT POVCAT15=POVCAT
              TOTSLF15=TOTSLF));
RUN;
DATA YR2X;
       SET YR2 (RENAME=(INSCOV16=INSCOV PERWT16F=PERWT POVCAT16=POVCAT
              TOTSLF16=TOTSLF));
RUN;
/*********************************************************************************
Create a pooled dataset with the two-year (2015,2016) datasets created above
Insurance coverage, last age, poverty category, as well as variables for
CLUSTER, STRATA, and WEIGHT are kept
Divide the PERWT by 2 because we are now dealing with datasets in 2 years
 - assign to the POOLWT variable

Create a FLAG if individuals are Between 26 and 30 years of AGE
(range inclusive), UNINSURED, and HIGH INCOME
Create LABELS to enhance readability for SUBPOP variable and TOTSLF
(Amount paid by self/family)
 ********************************************************************************/;
DATA POOL;
       LENGTH INSCOV AGELAST POVCAT VARSTR VARPSU 8;
       SET YR1X YR2X;
       POOLWT=PERWT/2;
       IF 26 LE AGELAST LE 30 AND POVCAT=5 AND INSCOV=3 THEN
              SUBPOP=1;
       ELSE
              SUBPOP=2;
       LABEL
              SUBPOP='POPULATION WITH AGE=26-30, UNINSURED WHOLE YEAR, AND HIGH INCOME'
              TOTSLF='TOTAL AMT PAID BY SELF/FAMILY';
RUN;
/*********************************************************************************
Check missing values on the combined dataset (2015 and 2016)
(N provides # of observations, while NMISS provides # of observations missing)
 ********************************************************************************/;
PROC MEANS DATA=POOL N NMISS;
       TITLE3 "CHECK MISSING VALUES ON THE COMBINED DATA";
RUN;
/*********************************************************************************
Display FREQ counts for both years combined (2015, 2016)
Display FREQ for SUBPOP (26-30, high income, uninsured) Flag
Display FREQ for SUBPOP (26-30, high income, uninsured) by AGE, POVCAT, and INSCOV
 ********************************************************************************/;
PROC FREQ DATA=POOL;
       TITLE3 'SUPPORTING CROSSTAB FOR THE CREATION OF THE SUBPOP FLAG';
       TABLES SUBPOP SUBPOP*AGELAST*POVCAT*INSCOV/ LIST MISSING;
       FORMAT AGELAST AGE.;
RUN;
/* Suppress the printing of output */
ODS EXCLUDE ALL;
/*********************************************************************************
Use the SURVEYMEANS PROC to create estimates for larger population based on the
design of our survey sample
OUTPUT is temporarily suppressed and later used in PRINT PROC to display results
CALCULATIONS show weighted estimated EXPENDITURES for SUBPOP
(AGE:26-30,Uninsured Whole Year,High Income) for combined years (2015,2016)
OUTPUT DOMAIN Result for Further processing in PROC PRINT statement
 ********************************************************************************/;
```

```
PROC SURVEYMEANS DATA=POOL NOBS MEAN STDERR;
       TITLE3 'WEIGHTED ESTIMATE ON TOTSLF FOR COMBINED DATA W/AGE=26-30,
       UNINSURED WHOLE YEAR, AND HIGH INCOME';
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT POOLWT;
       VAR TOTSLF;
       DOMAIN SUBPOP("1");
       ODS OUTPUT DOMAIN=work.domain_results;
RUN;
/* Unsuppress the printing of output */
ODS EXCLUDE NONE;
/*****************************************************************************
Print the results of the DOMAIN table from the dataset output above
NOOBS= SUPPRESS # OF OBSERVATIONS
SPLIT= SPLIT ON DESIGNATED FIELD IN COLUMN HEADINGS
Display WEIGHTED ESTIMATE ON TOTSLF FOR COMBINED DATA W/ AGE=26-30,
UNINSURED WHOLE YEAR, AND HIGH INCOME'
Display the estimated amount paid by SELF/FAMILY for SUBPOP for combined years
 *****************************************************************************/;
proc print data=work.domain_results noobs split='*';
       var VARLABEL N mean StdErr;
       label mean='Mean' StdErr='SE of Mean';
       format N Comma12. mean comma9.1 stderr 9.4;
       TITLE3 'WEIGHTED ESTIMATE ON TOTSLF FOR COMBINED DATA W/AGE=26-30, UNINSURED WHOLE YEAR, AND
HIGH INCOME';
run;
/* CLOSE ALL OUTPUT */
ODS _ALL_ CLOSE;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
PROC PRINTTO;
RUN;
```

Exercise 4.A Output

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE6.SAS: POOL MEPS DATA FILES FROM DIFFERENT YEARS (2015 and 2016)
Summary of H181 Variables

The CONTENTS Procedure

Variables in Creation Order

#	Variable	Type	Len	Label
1	DUPERSID	Char	8	PERSON ID (DUID + PID)
2	AGELAST	Num	4	PERSON'S AGE LAST TIME ELIGIBLE
3	POVCAT15	Num	4	FAMILY INC AS % OF POVERTY LINE - CATEGO
4	INSCOV15	Num	4	HEALTH INSURANCE COVERAGE INDICATOR 2015
5	TOTSLF15	Num	8	TOTAL AMT PAID BY SELF/FAMILY 15
6	PERWT15F	Num	8	FINAL PERSON WEIGHT, 2015
7	VARSTR	Num	5	VARIANCE ESTIMATION STRATUM - 2015
8	VARPSU	Num	4	VARIANCE ESTIMATION PSU - 2015

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE6.SAS: POOL MEPS DATA FILES FROM DIFFERENT YEARS (2015 and 2016)
UNWEIGHTED FREQUENCY FOR 2015 FY PERSONS WITH AGE 26-30

The FREQ Procedure

POVCAT15	INSCOV15	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1 POOR/NEGATIVE	1 ANY PRIVATE	83	3.61	83	3.61
1 POOR/NEGATIVE	2 PUBLIC ONLY	236	10.27	319	13.88
1 POOR/NEGATIVE	3 UNINSURED	145	6.31	464	20.19
2 NEAR POOR	1 ANY PRIVATE	38	1.65	502	21.85
2 NEAR POOR	2 PUBLIC ONLY	52	2.26	554	24.11
2 NEAR POOR	3 UNINSURED	37	1.61	591	25.72
3 LOW INCOME	1 ANY PRIVATE	182	7.92	773	33.64
3 LOW INCOME	2 PUBLIC ONLY	78	3.39	851	37.03

3	LOW INCOME	3	UNINSURED	94	4.09	945	41.12
4	MIDDLE INCOME	1	ANY PRIVATE	500	21.76	1445	62.88
4	MIDDLE INCOME	2	PUBLIC ONLY	90	3.92	1535	66.80
4	MIDDLE INCOME	3	UNINSURED	140	6.09	1675	72.89
5	HIGH INCOME	1	ANY PRIVATE	547	23.80	2222	96.69
5	HIGH INCOME	2	PUBLIC ONLY	29	1.26	2251	97.95
5	HIGH INCOME	3	UNINSURED	47	2.05	2298	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE6.SAS: POOL MEPS DATA FILES FROM DIFFERENT YEARS (2015 and 2016)
UNWEIGHTED FREQUENCY FOR 2016 FY PERSONS WITH AGE 26-30

The FREQ Procedure

POVCAT16	INSCOV16	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1 POOR/NEGATIVE	1 ANY PRIVATE	88	4.20	88	4.20
1 POOR/NEGATIVE	2 PUBLIC ONLY	208	9.93	296	14.13
1 POOR/NEGATIVE	3 UNINSURED	118	5.63	414	19.76
2 NEAR POOR	1 ANY PRIVATE	29	1.38	443	21.15
2 NEAR POOR	2 PUBLIC ONLY	44	2.10	487	23.25
2 NEAR POOR	3 UNINSURED	35	1.67	522	24.92
3 LOW INCOME	1 ANY PRIVATE	137	6.54	659	31.46
3 LOW INCOME	2 PUBLIC ONLY	84	4.01	743	35.47
3 LOW INCOME	3 UNINSURED	91	4.34	834	39.81
4 MIDDLE INCOME	1 ANY PRIVATE	500	23.87	1334	63.68
4 MIDDLE INCOME	2 PUBLIC ONLY	72	3.44	1406	67.11
4 MIDDLE INCOME	3 UNINSURED	115	5.49	1521	72.60
5 HIGH INCOME	1 ANY PRIVATE	490	23.39	2011	95.99
5 HIGH INCOME	2 PUBLIC ONLY	25	1.19	2036	97.18
5 HIGH INCOME	3 UNINSURED	59	2.82	2095	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE6.SAS: POOL MEPS DATA FILES FROM DIFFERENT YEARS (2015 and 2016)
CHECK MISSING VALUES ON THE COMBINED DATA

The MEANS Procedure

Variable	Label	N	N Miss
INSCOV	HEALTH INSURANCE COVERAGE INDICATOR 2015	67152	0
AGELAST	PERSON'S AGE LAST TIME ELIGIBLE	67152	0
POVCAT	FAMILY INC AS % OF POVERTY LINE - CATEGO	67152	0
VARSTR	VARIANCE ESTIMATION STRATUM - 2015	67152	0
VARPSU	VARIANCE ESTIMATION PSU - 2015	67152	0
TOTSLF	TOTAL AMT PAID BY SELF/FAMILY	67152	0
PERWT	FINAL PERSON WEIGHT, 2015	67152	0
POOLWT		67152	0
SUBPOP	POPULATION WITH AGE=26-30, UNINSURED WHOLE YEAR, AND HIGH INCOME	67152	0

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE6.SAS: POOL MEPS DATA FILES FROM DIFFERENT YEARS (2015 and 2016)
SUPPORTING CROSSTAB FOR THE CREATION OF THE SUBPOP FLAG

The FREQ Procedure

POPULATION WITH AGE=26-30, UNINSURED WHOLE YEAR, AND HIGH INCOME

SUBPOP	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	106	0.16	106	0.16
2	67046	99.84	67152	100.00

SUBPOP	AGELAST	POVCAT	INSCOV	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	26-30	5	3	106	0.16	106	0.16
2	0-25	1	1	846	1.26	952	1.42
2	0-25	1	2	6398	9.53	7350	10.95
2	0-25	1	3	528	0.79	7878	11.73
2	0-25	2	1	394	0.59	8272	12.32
2	0-25	2	2	1323	1.97	9595	14.29
2	0-25	2	3	168	0.25	9763	14.54

2	0-25	3	1	1451	2.16	11214	16.70
2	0-25	3	2	2593	3.86	13807	20.56
2	0-25	3	3	449	0.67	14256	21.23
2	0-25	4	1	4250	6.33	18506	27.56
2	0-25	4	2	1711	2.55	20217	30.11
2	0-25	4	3	523	0.78	20740	30.89
2	0-25	5	1	4021	5.99	24761	36.87
2	0-25	5	2	242	0.36	25003	37.23
2	0-25	5	3	190	0.28	25193	37.52
2	26-30	1	1	171	0.25	25364	37.77
2	26-30	1	2	444	0.66	25808	38.43
2	26-30	1	3	263	0.39	26071	38.82
2	26-30	2	1	67	0.10	26138	38.92
2	26-30	2	2	96	0.14	26234	39.07
2	26-30	2	3	72	0.11	26306	39.17
2	26-30	3	1	319	0.48	26625	39.65
2	26-30	3	2	162	0.24	26787	39.89
2	26-30	3	3	185	0.28	26972	40.17
2	26-30	4	1	1000	1.49	27972	41.65
2	26-30	4	2	162	0.24	28134	41.90
2	26-30	4	3	255	0.38	28389	42.28
2	26-30	5	1	1037	1.54	29426	43.82
2	26-30	5	2	54	0.08	29480	43.90
2	31+	1	1	1125	1.68	30605	45.58
2	31+	1	2	3690	5.49	34295	51.07
2	31+	1	3	1364	2.03	35659	53.10
2	31+	2	1	606	0.90	36265	54.00
2	31+	2	2	1047	1.56	37312	55.56
2	31+	2	3	426	0.63	37738	56.20
2	31+	3	1	2521	3.75	40259	59.95
2	31+	3	2	2268	3.38	42527	63.33
2	31+	3	3	1072	1.60	43599	64.93
2	31+	4	1	7508	11.18	51107	76.11
2	31+	4	2	2291	3.41	53398	79.52
2	31+	4	3	1215	1.81	54613	81.33
2	31+	5	1	10842	16.15	65455	97.47
2	31+	5	2	1206	1.80	66661	99.27
2	31+	5	3	491	0.73	67152	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE6.SAS: POOL MEPS DATA FILES FROM DIFFERENT YEARS (2015 and 2016)
WEIGHTED ESTIMATE ON TOTSLF FOR COMBINED DATA W/AGE=26-30, UNINSURED WHOLE YEAR, AND HIGH INCOME

Variable Label	N	Mean	SE of Mean
TOTAL AMT PAID BY SELF/FAMILY	106	241.5	80.8841

Exercise 4.B Script

```
/*************************************************************************
DESCRIPTION:  THIS PROGRAM ILLUSTRATES HOW TO POOL MEPS LONGITUDINAL DATA
FILES FROM DIFFERENT PANELS
THE EXAMPLE USED IS PANELS 17-19 POPULATION AGE 26-30 WHO ARE
UNINSURED BUT HAVE HIGH INCOME IN THE FIRST YEAR
DATA FROM PANELS 17, 18, AND 19 ARE POOLED.
INPUT FILE:     (1) C:\MEPS\SAS\DATA\H183.SAS7BDAT (PANEL 19 LONGITUDINAL FILE)
(2) C:\MEPS\SAS\DATA\H172.SAS7BDAT (PANEL 18 LONGITUDINAL FILE)
(3) C:\MEPS\SAS\DATA\H164.SAS7BDAT (PANEL 17 LONGITUDINAL FILE)
 *************************************************************************/;
OPTIONS NODATE NONUMBER PAGESIZE=200 FORMCHAR="|----|+|---+=|-/\<>*";
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=4b;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
/* Suppress the date and time in the output */
OPTIONS NODATE;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 'EXERCISE 4B SAS: POOL MEPS DATA FILES
FROM DIFFERENT PANELS (PANELS 17, 18, 19)';
/* Create format variables to display Poverty Category,
Insurance status, and Age Subpopulation variables */
```

```
PROC FORMAT;
       VALUE POVCAT 1='1 POOR/NEGATIVE' 2='2 NEAR POOR' 3='3 LOW INCOME'
              4='4 MIDDLE INCOME' 5='5 HIGH INCOME';
       VALUE INSF-1='-1 INAPPLICABLE' 1='1 ANY PRIVATE' 2='2 PUBLIC ONLY'
              3='3 UNINSURED';
       VALUE AGE -1='-1 INAPPLICABLE' 26-30='26-30' 0-25, 31-HIGH='0-25, 31+';
run;
/*************************************************************************
Select subset of variables from main data table
Display summary statistics for data file
 *************************************************************************/;
DATA SUMMARYSET (KEEP=DUPERSID INSCOVY1 INSCOVY2 LONGWT VARSTR VARPSU POVCATY1
              AGEY1X PANEL);
       SET CDATA.H164;
RUN;
/* Select only the position table in the output */
ODS Select Position;
/* Display summary statistics: File #2 */
PROC CONTENTS DATA=WORK.SUMMARYSET VARNUM;
       Title3 'Summary of Datafile Variables for LONGITUDINAL Files';
RUN;
/*************************************************************************
Create a pooled dataset based off 3 data files
Insurance coverage, age, poverty category, as well as variables for
CLUSTER, STRATA, and WEIGHT are kept
Divide the LONGWT VARIABLE by 3 because we are now dealing with datasets in 3 years
 - assign to the POOLWT variable
Create a SUBPOP FLAG to denote Persons (26-30, high income, uninsured) IN FIRST
YEAR
 *************************************************************************/;
DATA POOL;
       LENGTH INSCOVY1 INSCOVY2 PANEL AGEY1X POVCATY1 VARSTR VARPSU 8;
       SET CDATA.H164 (KEEP=DUPERSID INSCOVY1 INSCOVY2 LONGWT VARSTR VARPSU POVCATY1
              AGEY1X PANEL) CDATA.H172 (KEEP=DUPERSID INSCOVY1 INSCOVY2 LONGWT VARSTR
              VARPSU POVCATY1 AGEY1X PANEL) CDATA.H183 (KEEP=DUPERSID INSCOVY1 INSCOVY2
              LONGWT VARSTR VARPSU POVCATY1 AGEY1X PANEL);
       POOLWT=LONGWT/3;
       IF INSCOVY1=3 AND 26 LE AGEY1X LE 30 AND POVCATY1=5 THEN
              SUBPOP=1;
       ELSE
              SUBPOP=2;
       LABEL
              SUBPOP='POPULATION WITH AGE=26-30, UNINSURED, AND HIGH INCOME IN FIRST YEAR'
              INSCOVY2="HEALTH INSURANCE COVERAGE INDICATOR IN YEAR 2";
RUN;
/* Check missing values on the combined dataset (2015 and 2016)
(N provides # of observations, while NMISS provides
# of observations missing) */
PROC MEANS DATA=POOL N NMISS;
       TITLE3 "CHECK MISSING VALUES ON THE COMBINED DATA";
RUN;
/*************************************************************************
Create FREQ count for SUBPOP Flag (26-30, high income, uninsured)
Create FREQ display for SUBPOP FLAG (26-30, high income, uninsured) By Survey Panel
Create FREQ display for SUBPOP FLAG (26-30, high income, uninsured)
Insurance Coverage Indicator, Age as of 12/31/12, and Poverty Category in first year
 *************************************************************************/;
PROC FREQ DATA=POOL;
       TABLES SUBPOP SUBPOP*PANEL SUBPOP*INSCOVY1*AGEY1X*POVCATY1/LIST MISSING;
       FORMAT AGEY1X AGE. POVCATY1 POVCAT. INSCOVY1 INSF.;
       TITLE3 'SUPPORTING CROSSTAB FOR THE CREATION OF THE SUBPOP FLAG';
RUN;
/* Turn all graphs off */
ODS GRAPHICS OFF;
/* Suppress the printing of output */
ODS EXCLUDE ALL;
/*************************************************************************
Estimate Insurance Coverage Indicator for 2nd year of SUBPOP
(26-30, high income, uninsured)
```

```
Calculate the estimated proportion of Insurance Coverage Indicator for
SUBPOP in Year 2
Pipe DOMAIN output to temporary dataset for further processing by
PROC PRINT procedure
 ********************************************************************************/;
PROC SURVEYMEANS DATA=POOL NOBS MEAN STDERR;
       STRATUM VARSTR;
       CLUSTER VARPSU;
       WEIGHT POOLWT;
       VAR INSCOVY2;
       CLASS INSCOVY2;
       FORMAT INSCOVY2 INSF.;
       DOMAIN SUBPOP('1');
       ODS OUTPUT DOMAIN=work.domain_results;
       TITLE3 'INSURANCE STATUS IN THE SECOND YEAR FOR THOSE W/ AGE=26-30,
       UNINSURED WHOLE YEAR, AND HIGH INCOME IN THE FIRST YEAR';
RUN;
/* Unsuppress the printing of output */
ODS EXCLUDE NONE;
/*********************************************************************************
Print the results of the DOMAIN table from the dataset output above
NOOBS= SUPPRESS # OF OBSERVATIONS
SPLIT= SPLIT ON DESIGNATED FIELD IN COLUMN HEADINGS
VAR specifies the designated variables to display
LABEL and FORMAT provide appropriate display options
Display Estimated Insurance Coverage Status for SUBPOP in 2nd Year
 ********************************************************************************/;
proc print data=work.domain_results noobs split='*';
       var VARLEVEL N mean StdErr;
       label mean='Proportion' StdErr='SE of Proportion';
       format N Comma12. mean comma9.3 stderr 9.6;
       TITLE3 'INSURANCE STATUS IN THE SECOND YEAR FOR THOSE W/ AGE=26-30,
       UNINSURED WHOLE YEAR, AND HIGH INCOME IN THE FIRST YEAR';
run;
/* Close all the output */
ODS _ALL_ CLOSE;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
PROC PRINTTO;
RUN;
```

Exercise 4.B Output

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 4B SAS: POOL MEPS DATA FILES FROM DIFFERENT PANELS (PANELS 17, 18, 19)
Summary of Datafile Variables for LONGITUDINAL Files

The CONTENTS Procedure

Variables in Creation Order

#	Variable	Type	Len	Label
1	DUPERSID	Char	8	PERSON ID (DUID + PID)
2	PANEL	Num	8	PANEL NUMBER
3	AGEY1X	Num	8	AGE AS OF 12/31/12 (EDITED/IMPUTED)
4	POVCATY1	Num	8	FAMLY INC AS % OF POVERTY LINE-CATEGO 12
5	INSCOVY1	Num	8	HEALTH INSURANCE COVERAGE INDICATOR 12
6	INSCOVY2	Num	4	HEALTH INSURANCE COVERAGE INDICATOR 13
7	LONGWT	Num	8	LONGITUDINAL WEIGHT - PANEL 17
8	VARPSU	Num	8	VARIANCE ESTIMATION PSU
9	VARSTR	Num	8	VARIANCE ESTIMATION STRATUM

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 4B SAS: POOL MEPS DATA FILES FROM DIFFERENT PANELS (PANELS 17, 18, 19)
CHECK MISSING VALUES ON THE COMBINED DATA

The MEANS Procedure

Variable	Label	N	N Miss
INSCOVY1	HEALTH INSURANCE COVERAGE INDICATOR 12	50535	0
INSCOVY2	HEALTH INSURANCE COVERAGE INDICATOR IN YEAR 2	50535	0
PANEL	PANEL NUMBER	50535	0
AGEY1X	AGE AS OF 12/31/12 (EDITED/IMPUTED)	50535	0

POVCATY1	FAMLY INC AS % OF POVERTY LINE-CATEGO 12	50535	0
VARSTR	VARIANCE ESTIMATION STRATUM	50535	0
VARPSU	VARIANCE ESTIMATION PSU	50535	0
LONGWT	LONGITUDINAL WEIGHT - PANEL 17	50535	0
POOLWT		50535	0
SUBPOP	POPULATION WITH AGE=26-30, UNINSURED, AND HIGH INCOME IN FIRST YEAR	50535	0

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 4B SAS: POOL MEPS DATA FILES FROM DIFFERENT PANELS (PANELS 17, 18, 19)
SUPPORTING CROSSTAB FOR THE CREATION OF THE SUBPOP FLAG

The FREQ Procedure

POPULATION WITH AGE=26-30, UNINSURED, AND HIGH INCOME IN FIRST YEAR

SUBPOP	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	78	0.15	78	0.15
2	50457	99.85	50535	100.00

SUBPOP	PANEL	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	17	41	0.08	41	0.08
1	18	22	0.04	63	0.12
1	19	15	0.03	78	0.15
2	17	17882	35.39	17960	35.54
2	18	16692	33.03	34652	68.57
2	19	15883	31.43	50535	100.00

SUBPOP	INSCOVY1	AGEY1X	POVCATY1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	3 UNINSURED	26-30	5 HIGH INCOME	78	0.15	78	0.15
2	-1 INAPPLICABLE	-1 INAPPLICABLE	-1	647	1.28	725	1.43
2	1 ANY PRIVATE	-1 INAPPLICABLE	1 POOR/NEGATIVE	25	0.05	750	1.48
2	1 ANY PRIVATE	-1 INAPPLICABLE	2 NEAR POOR	10	0.02	760	1.50
2	1 ANY PRIVATE	-1 INAPPLICABLE	3 LOW INCOME	21	0.04	781	1.55
2	1 ANY PRIVATE	-1 INAPPLICABLE	4 MIDDLE INCOME	32	0.06	813	1.61
2	1 ANY PRIVATE	-1 INAPPLICABLE	5 HIGH INCOME	34	0.07	847	1.68
2	1 ANY PRIVATE	0-25, 31+	1 POOR/NEGATIVE	1361	2.69	2208	4.37
2	1 ANY PRIVATE	0-25, 31+	2 NEAR POOR	685	1.36	2893	5.72
2	1 ANY PRIVATE	0-25, 31+	3 LOW INCOME	3048	6.03	5941	11.76
2	1 ANY PRIVATE	0-25, 31+	4 MIDDLE INCOME	8270	16.36	14211	28.12
2	1 ANY PRIVATE	0-25, 31+	5 HIGH INCOME	9717	19.23	23928	47.35
2	1 ANY PRIVATE	26-30	1 POOR/NEGATIVE	107	0.21	24035	47.56
2	1 ANY PRIVATE	26-30	2 NEAR POOR	63	0.12	24098	47.69
2	1 ANY PRIVATE	26-30	3 LOW INCOME	273	0.54	24371	48.23
2	1 ANY PRIVATE	26-30	4 MIDDLE INCOME	668	1.32	25039	49.55
2	1 ANY PRIVATE	26-30	5 HIGH INCOME	656	1.30	25695	50.85
2	2 PUBLIC ONLY	-1 INAPPLICABLE	1 POOR/NEGATIVE	59	0.12	25754	50.96
2	2 PUBLIC ONLY	-1 INAPPLICABLE	2 NEAR POOR	17	0.03	25771	51.00
2	2 PUBLIC ONLY	-1 INAPPLICABLE	3 LOW INCOME	30	0.06	25801	51.06
2	2 PUBLIC ONLY	-1 INAPPLICABLE	4 MIDDLE INCOME	43	0.09	25844	51.14
2	2 PUBLIC ONLY	-1 INAPPLICABLE	5 HIGH INCOME	22	0.04	25866	51.18
2	2 PUBLIC ONLY	0-25, 31+	1 POOR/NEGATIVE	7725	15.29	33591	66.47
2	2 PUBLIC ONLY	0-25, 31+	2 NEAR POOR	1832	3.63	35423	70.10
2	2 PUBLIC ONLY	0-25, 31+	3 LOW INCOME	3132	6.20	38555	76.29
2	2 PUBLIC ONLY	0-25, 31+	4 MIDDLE INCOME	2385	4.72	40940	81.01
2	2 PUBLIC ONLY	0-25, 31+	5 HIGH INCOME	865	1.71	41805	82.72
2	2 PUBLIC ONLY	26-30	1 POOR/NEGATIVE	320	0.63	42125	83.36
2	2 PUBLIC ONLY	26-30	2 NEAR POOR	71	0.14	42196	83.50
2	2 PUBLIC ONLY	26-30	3 LOW INCOME	120	0.24	42316	83.74
2	2 PUBLIC ONLY	26-30	4 MIDDLE INCOME	88	0.17	42404	83.91
2	2 PUBLIC ONLY	26-30	5 HIGH INCOME	23	0.05	42427	83.96
2	3 UNINSURED	-1 INAPPLICABLE	1 POOR/NEGATIVE	33	0.07	42460	84.02
2	3 UNINSURED	-1 INAPPLICABLE	2 NEAR POOR	7	0.01	42467	84.03
2	3 UNINSURED	-1 INAPPLICABLE	3 LOW INCOME	12	0.02	42479	84.06
2	3 UNINSURED	-1 INAPPLICABLE	4 MIDDLE INCOME	14	0.03	42493	84.09
2	3 UNINSURED	-1 INAPPLICABLE	5 HIGH INCOME	7	0.01	42500	84.10
2	3 UNINSURED	0-25, 31+	1 POOR/NEGATIVE	2194	4.34	44694	88.44
2	3 UNINSURED	0-25, 31+	2 NEAR POOR	674	1.33	45368	89.78
2	3 UNINSURED	0-25, 31+	3 LOW INCOME	1735	3.43	47103	93.21
2	3 UNINSURED	0-25, 31+	4 MIDDLE INCOME	1842	3.64	48945	96.85
2	3 UNINSURED	0-25, 31+	5 HIGH INCOME	598	1.18	49543	98.04
2	3 UNINSURED	26-30	1 POOR/NEGATIVE	342	0.68	49885	98.71
2	3 UNINSURED	26-30	2 NEAR POOR	130	0.26	50015	98.97
2	3 UNINSURED	26-30	3 LOW INCOME	240	0.47	50255	99.45
2	3 UNINSURED	26-30	4 MIDDLE INCOME	280	0.55	50535	100.00

Variable Level	N	Proportion	SE of Proportion
-1 INAPPLICABLE	0	0.000	0.000000
1 ANY PRIVATE	15	0.233	0.066438
2 PUBLIC ONLY	4	0.035	0.019375
3 UNINSURED	59	0.732	0.067309

Exercise 5.A Script

```
/*************************************************************************
DESCRIPTION:  THIS PROGRAM ILLUSTRATES HOW TO CONSTRUCT FAMILY-LEVEL VARIABLES
FROM PERSON-LEVEL DATA
THERE ARE TWO DEFINITIONS OF FAMILY UNIT IN MEPS.
1) CPS FAMILY:  ID IS DUID + CPSFAMID.  CORRESPONDING WEIGHT IS FAMWT15C.
2) MEPS FAMILY: ID IS DUID + FAMIDYR.   CORRESPONDING WEIGHT IS FAMWT15F.
THE CPS FAMILY IS USED IN THIS EXERCISE.
INPUT FILE:   C:\MEPS\SAS\DATA\H181.SAS7BDAT (2015 FY PUF DATA)
 *************************************************************************/;
/* Update log and print output information based on exercise # */
%LET folderPath=/folders/myfolders/sasuser.v94/MEPS/GitHub/MEPS/SAS/workshop_exercises/exercise_;
%LET exerciseNumber=5a;
%LET logFolderOutput=&folderPath.&exerciseNumber/my-output/;
OPTIONS LINESIZE=160 PAGESIZE=79 NODATE FORMCHAR="|----|+|---+=|-/\<>*"
      PAGENO=1;
FILENAME MYLOG "&logFolderOutput\Exercise&exerciseNumber._log.TXT";
FILENAME MYPRINT "&logFolderOutput\Exercise&exerciseNumber._OUTPUT.TXT";
PROC PRINTTO LOG=MYLOG PRINT=MYPRINT NEW;
RUN;
TITLE1 '2018 AHRQ MEPS DATA USERS WORKSHOP';
TITLE2 "EXERCISE 5a SAS: CALCULATE FAMILY-LEVEL ESTIMATES";
/* Summary statistics table 2 */
DATA SUMMARYSET (KEEP=DUID PID AGE15X SEX CPSFAMID FAMWT15C FAMIDYR FAMWT15F);
      SET CDATA.H181;
RUN;
/* Select only the position table in the output */
ODS Select Position;
/* Display summary statistics: File #2 */
PROC CONTENTS DATA=WORK.SUMMARYSET VARNUM;
      Title3 'Summary of H181 Variables';
RUN;
/* Select all output once again */
ODS Select All;
/*************************************************************************
The CPSFAMID represents a redefinition of MEPS families into
families defined by the Current Population Survey (CPS).
Some of the distinctions between CPS-and MEPS-defined families are that MEPS
families include and CPS families do not include non-married partners,
foster children, and in-laws. These persons are considered as members of separate
families for CPS-like families. CPS-like families are defined so a poverty status
classification variable consistent with established definitions of poverty can be
assigned to the CPS-like families and used for weight poststratification purposes.
FAMWT15C can be used to make estimates for the cross-section of families in the
U.S. civilian, noninstitutionalized population on December 31, 2015
where families are identified based on the CPS definition of a family unit.
The annualized family ID letter, FAMIDYR, identifies eligible members of the
eligible annualized families within a DU.
FAMWT15F can be used to make estimates for the cross-section of families in
the U.S. civilian noninstitutionalized population on December 31, 2015
where families are identified based on the MEPS definition of a family unit.
Each MEPS family unit is uniquely identified by the combination of the
variables DUID and FAMIDYR while each CPS family unit
is uniquely identified by the combination of the variables DUID and CPSFAMID.
 *************************************************************************/;
/* PROC PRINT FIRST 30 OBSERVATIONS AND SPECIFIED VARIABLES */
PROC PRINT DATA=CDATA.H181 (OBS=20);
```

```
      VAR PID AGE15X SEX CPSFAMID FAMWT15C FAMIDYR FAMWT15F;
      TITLE3 "SAMPLE DUMP FOR FAMILY IDS";
      BY DUID;
RUN;
/*****************************************************************************
Sort data by DUID and CPSFAMID to represent families defined by the Current
Population Survey (CPS)
Include ID variables, necessary weighting variables, total expenditures, and total
income in new dataset named PERS
 *****************************************************************************/;
PROC SORT DATA=CDATA.H181 (KEEP=DUPERSID DUID CPSFAMID FAMWT15C VARSTR VARPSU
            TOTSLF15 TTLP15X) OUT=PERS;
      BY DUID CPSFAMID;
RUN;
/*****************************************************************************
Create a dataset PERS2
Keep Dwelling Unit ID, CPS Family Id
Create variables, FAMSIZE (size), FAMOOP (out-of-pocket), and FAMINC (income)
Define values for newly created family variables based on a loop of the CPSFAMID
 *****************************************************************************/;
DATA PERS2 FAM (KEEP=DUID CPSFAMID FAMSIZE FAMOOP FAMINC);
      SET PERS;
      BY DUID CPSFAMID;
      LABEL FAMSIZE='# OF PERSONS PER CPS FAMILY'
            FAMOOP='TOTAL OUT-OF-POCKET EXP (TOTSLF15) PER CPS FAMILY'
            FAMINC='TOTAL INCOME (TTLP15X) PER CPS FAMILY';
      IF FIRST.CPSFAMID THEN
            DO;
                  FAMSIZE=0;
                  FAMOOP=0;
                  FAMINC=0;
            END;
      FAMSIZE + 1;
      FAMOOP  + TOTSLF15;
      FAMINC  + TTLP15X;
      OUTPUT PERS2;
      IF LAST.CPSFAMID THEN
            OUTPUT FAM;
RUN;
/*****************************************************************************
Print the PERS2 dataset created above (first 20 observations)
Print represents families defined by the Current Population Survey
 *****************************************************************************/;
PROC PRINT DATA=PERS2 (OBS=20);
      BY DUID CPSFAMID;
      TITLE3 "A SAMPLE DUMP TO CHECK THE CREATION OF THE FAMILY-LEVEL VARIABLES";
RUN;
/*****************************************************************************
ADD WEIGHT, VARSTR, AND VARPSU TO THE FAMILY-LEVEL ANALYTIC DATA
Sort the PERS dataset where FAMWT15C > 0
De-duplicating the results and keeping variables to use for PROC SURVEYMEANS statement
 *****************************************************************************/;
PROC SORT DATA=PERS (WHERE=(FAMWT15C>0)) OUT=FAMWT (KEEP=DUID CPSFAMID FAMWT15C
            VARSTR VARPSU) NODUPKEY;
      BY DUID CPSFAMID;
RUN;
/*****************************************************************************
Merge FAM data with FAMWT data created above By Dwelling-Unit ID and CPS Family ID
FAM2 is CPS families with WEIGHT variables for PROC SURVEYMEANS statement
 *****************************************************************************/;
DATA FAM2;
      MERGE FAM (IN=AA) FAMWT (IN=BB);
      BY DUID CPSFAMID;
      IF AA AND BB;
RUN;
/* Turn off graphics */
ods graphics off;
/*****************************************************************************
Provide CPS Family-Level Estimates on Family Size, CPS Family OUT-OF-POCKET EXP,
```

```
and CPS Family Income
 ********************************************************************************/;
PROC SURVEYMEANS DATA=FAM2 NOBS SUMWGT MEAN STDERR;
       STRATA VARSTR;
       CLUSTER VARPSU;
       WEIGHT FAMWT15C;
       TITLE3 "CPS FAMILY-LEVEL ESTIMATES ON FAMILY SIZE, OUT-OF-POCKET EXP,
AND INCOME, 2015";
       VAR FAMSIZE FAMOOP FAMINC;
RUN;
/* Close all the output */
ODS _ALL_ CLOSE;
/* THE PROC PRINTTO null step is required to close the PROC PRINTTO */
PROC PRINTTO;
RUN;
```

Exercise 5.A Output

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 5a SAS: CALCULATE FAMILY-LEVEL ESTIMATES
Summary of H181 Variables

The CONTENTS Procedure

Variables in Creation Order

#	Variable	Type	Len	Label
1	DUID	Num	6	DWELLING UNIT ID
2	PID	Num	4	PERSON NUMBER
3	FAMIDYR	Char	2	ANNUAL FAMILY IDENTIFIER
4	CPSFAMID	Char	2	CPSFAMID
5	AGE15X	Num	4	AGE AS OF 12/31/15(EDITED/IMPUTED)
6	SEX	Num	4	SEX
7	FAMWT15F	Num	8	FINAL FAMILY WEIGHT, 2015
8	FAMWT15C	Num	8	POV ADJ FAMILY WGT-CPS FAM ON 12/31/15

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 5a SAS: CALCULATE FAMILY-LEVEL ESTIMATES
SAMPLE DUMP FOR FAMILY IDS

DWELLING UNIT ID=60001

Obs	PID	AGE15X	SEX	CPSFAMID	FAMWT15C	FAMIDYR	FAMWT15F
1	101	53	1	A	22983.49	A	22983.49
2	102	56	2	B	14663.34	A	22983.49
3	103	30	1	B	14663.34	A	22983.49
4	104	23	2	C	11200.74	A	22983.49
5	105	3	1	C	11200.74	A	22983.49

DWELLING UNIT ID=60002

Obs	PID	AGE15X	SEX	CPSFAMID	FAMWT15C	FAMIDYR	FAMWT15F
6	101	27	1	A	7240.72	A	7240.72
7	102	25	2	A	7240.72	A	7240.72
8	103	4	1	A	7240.72	A	7240.72
9	104	1	2	A	7240.72	A	7240.72

DWELLING UNIT ID=60003

Obs	PID	AGE15X	SEX	CPSFAMID	FAMWT15C	FAMIDYR	FAMWT15F
10	101	45	2	A	11240.31	A	11240.31
11	102	49	1	A	11240.31	A	11240.31
12	103	12	1	A	11240.31	A	11240.31
13	104	10	2	A	11240.31	A	11240.31

DWELLING UNIT ID=60004

Obs	PID	AGE15X	SEX	CPSFAMID	FAMWT15C	FAMIDYR	FAMWT15F
14	101	27	2	A	4105.35	A	4105.35
15	102	54	2	A	4105.35	A	4105.35
16	103	61	1	A	4105.35	A	4105.35

DWELLING UNIT ID=60009

Obs	PID	AGE15X	SEX	CPSFAMID	FAMWT15C	FAMIDYR	FAMWT15F
17	101	32	1	A	2564.82	A	2564.82

DWELLING UNIT ID=60010

Obs	PID	AGE15X	SEX	CPSFAMID	FAMWT15C	FAMIDYR	FAMWT15F
18	101	53	2	A	5736.58	A	5736.58
19	103	27	1	A	5736.58	A	5736.58
20	106	60	1	A	5736.58	A	5736.58

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 5a SAS: CALCULATE FAMILY-LEVEL ESTIMATES
A SAMPLE DUMP TO CHECK THE CREATION OF THE FAMILY-LEVEL VARIABLES

DWELLING UNIT ID=60001 CPSFAMID=A

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
1	60001101	11390	375	22983.49	1019	1	1	375	11390

DWELLING UNIT ID=60001 CPSFAMID=B

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
2	60001102	11390	11	14663.34	1019	1	1	11	11390
3	60001103	18000	0	14663.34	1019	1	2	11	29390

DWELLING UNIT ID=60001 CPSFAMID=C

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
4	60001104	18000	22	11200.74	1019	1	1	22	18000
5	60001105	385	77	11200.74	1019	1	2	99	18385

DWELLING UNIT ID=60002 CPSFAMID=A

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
6	60002101	33300	0	7240.72	1135	3	1	0	33300
7	60002102	3700	0	7240.72	1135	3	2	0	37000
8	60002103	0	153	7240.72	1135	3	3	153	37000
9	60002104	0	0	7240.72	1135	3	4	153	37000

DWELLING UNIT ID=60003 CPSFAMID=A

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
10	60003101	85000	57	11240.31	1054	1	1	57	85000
11	60003102	85000	90	11240.31	1054	1	2	147	170000
12	60003103	0	140	11240.31	1054	1	3	287	170000
13	60003104	0	60	11240.31	1054	1	4	347	170000

DWELLING UNIT ID=60004 CPSFAMID=A

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
14	60004101	29000	0	4105.35	1091	2	1	0	29000
15	60004102	35818	0	4105.35	1091	2	2	0	64818
16	60004103	700	602	4105.35	1091	2	3	602	65518

DWELLING UNIT ID=60009 CPSFAMID=A

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
17	60009101	24000	75	2564.82	1065	3	1	75	24000

---------- DWELLING UNIT ID=60010 CPSFAMID=A ----------

Obs	DUPERSID	TTLP15X	TOTSLF15	FAMWT15C	VARSTR	VARPSU	FAMSIZE	FAMOOP	FAMINC
18	60010101	0	0	5736.58	1036	2	1	0	0
19	60010103	10080	0	5736.58	1036	2	2	0	10080
20	60010106	24000	2	5736.58	1036	2	3	2	34080

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE 5a SAS: CALCULATE FAMILY-LEVEL ESTIMATES
CPS FAMILY-LEVEL ESTIMATES ON FAMILY SIZE, OUT-OF-POCKET EXP, AND INCOME, 2015

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	14522
Sum of Weights	141808362

Statistics

Variable	Label	N	Sum of Weights	Mean	Std Error of Mean
FAMSIZE	# OF PERSONS PER CPS FAMILY	14522	141808362	2.205817	0.019593
FAMOOP	TOTAL OUT-OF-POCKET EXP (TOTSLF15) PER CPS FAMILY	14522	141808362	1328.594367	32.955419
FAMINC	TOTAL INCOME (TTLP15X) PER CPS FAMILY	14522	141808362	64940	953.000134

Exercise 5.B Output

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE7.SAS: HOW TO CONSTRUCT INSURANCE STATUS VARIABLES, USING FY 15 DATA
Condensed list of variables

The DATASETS Procedure

Variables in Creation Order

#	Variable	Type	Len	Label
1	DUID	Num	6	DWELLING UNIT ID
2	PID	Num	4	PERSON NUMBER
3	DUPERSID	Char	8	PERSON ID (DUID + PID)
4	TRIJA15X	Num	4	COVERED BY TRICARE/CHAMPVA IN JAN15 (ED)
5	TRIFE15X	Num	4	COVERED BY TRICARE/CHAMPVA IN FEB15 (ED)
6	TRIMA15X	Num	4	COVERED BY TRICARE/CHAMPVA IN MAR15 (ED)
7	TRIAP15X	Num	4	COVERED BY TRICARE/CHAMPVA IN APR15(ED)
8	TRIMY15X	Num	4	COVERED BY TRICARE/CHAMPVA IN MAY15 (ED)
9	TRIJU15X	Num	4	COVERED BY TRICARE/CHAMPVA IN JUN15 (ED)
10	TRIJL15X	Num	4	COVERED BY TRICARE/CHAMPVA IN JUL15 (ED)
11	TRIAU15X	Num	4	COVERED BY TRICARE/CHAMPVA IN AUG15 (ED)
12	TRISE15X	Num	4	COVERED BY TRICARE/CHAMPVA IN SEP15 (ED)
13	TRIOC15X	Num	4	COVERED BY TRICARE/CHAMPVA IN OCT15 (ED)
14	TRINO15X	Num	4	COVERED BY TRICARE/CHAMPVA IN NOV15 (ED)
15	TRIDE15X	Num	4	COVERED BY TRICARE/CHAMPVA IN DEC15 (ED)
16	MCRJA15X	Num	4	COVERED BY MEDICARE IN JAN15 (ED)
17	MCRFE15X	Num	4	COVERED BY MEDICARE IN FEB15 (ED)
18	MCRMA15X	Num	4	COVERED BY MEDICARE IN MAR15 (ED)
19	MCRAP15X	Num	4	COVERED BY MEDICARE IN APR15 (ED)
20	MCRMY15X	Num	4	COVERED BY MEDICARE IN MAY15 (ED)
21	MCRJU15X	Num	4	COVERED BY MEDICARE IN JUN15 (ED)
22	MCRJL15X	Num	4	COVERED BY MEDICARE IN JUL15 (ED)
23	MCRAU15X	Num	4	COVERED BY MEDICARE IN AUG15 (ED)
24	MCRSE15X	Num	4	COVERED BY MEDICARE IN SEP15 (ED)
25	MCROC15X	Num	4	COVERED BY MEDICARE IN OCT15 (ED)
26	MCRNO15X	Num	4	COVERED BY MEDICARE IN NOV15 (ED)
27	MCRDE15X	Num	4	COVERED BY MEDICARE IN DEC15 (ED)
28	MCDJA15X	Num	4	COV BY MEDICAID OR SCHIP IN JAN15 (ED)
29	MCDFE15X	Num	4	COV BY MEDICAID OR SCHIP IN FEB15 (ED)
30	MCDMA15X	Num	4	COV BY MEDICAID OR SCHIP IN MAR15 (ED)
31	MCDAP15X	Num	4	COV BY MEDICAID OR SCHIP IN APR15 (ED)
32	MCDMY15X	Num	4	COV BY MEDICAID OR SCHIP IN MAY15 (ED)

33	MCDJU15X	Num	4	COV BY MEDICAID OR SCHIP IN JUN15 (ED)
34	MCDJL15X	Num	4	COV BY MEDICAID OR SCHIP IN JUL15 (ED)
35	MCDAU15X	Num	4	COV BY MEDICAID OR SCHIP IN AUG15 (ED)
36	MCDSE15X	Num	4	COV BY MEDICAID OR SCHIP IN SEP15 (ED)
37	MCDOC15X	Num	4	COV BY MEDICAID OR SCHIP IN OCT15 (ED)
38	MCDNO15X	Num	4	COV BY MEDICAID OR SCHIP IN NOV15 (ED)
39	MCDDE15X	Num	4	COV BY MEDICAID OR SCHIP IN DEC15 (ED)
40	PEGJA15	Num	4	COVERED BY EMPL UNION INS IN JAN15
41	PEGFE15	Num	4	COVERED BY EMPL UNION INS IN FEB15
42	PEGMA15	Num	4	COVERED BY EMPL UNION INS IN MAR15
43	PEGAP15	Num	4	COVERED BY EMPL UNION INS IN APR15
44	PEGMY15	Num	4	COVERED BY EMPL UNION INS IN MAY15
45	PEGJU15	Num	4	COVERED BY EMPL UNION INS IN JUN15
46	PEGJL15	Num	4	COVERED BY EMPL UNION INS IN JUL15
47	PEGAU15	Num	4	COVERED BY EMPL UNION INS IN AUG15
48	PEGSE15	Num	4	COVERED BY EMPL UNION INS IN SEP15
49	PEGOC15	Num	4	COVERED BY EMPL UNION INS IN OCT15
50	PEGNO15	Num	4	COVERED BY EMPL UNION INS IN NOV15
51	PEGDE15	Num	4	COVERED BY EMPL UNION INS IN DEC15
52	PDKJA15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) JAN15
53	PDKFE15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) FEB15
54	PDKMA15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) MAR15
55	PDKAP15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) APR15
56	PDKMY15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) MAY15
57	PDKJU15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) JUN15
58	PDKJL15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) JUL15
59	PDKAU15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) AUG15
60	PDKSE15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) SEP15
61	PDKOC15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) OCT15
62	PDKNO15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) NOV15
63	PDKDE15	Num	4	COVR BY PRIV INS (SOURCE UNKNWN) DEC15
64	PNGJA15	Num	4	COVERED BY NONGROUP INS IN JAN15
65	PNGFE15	Num	4	COVERED BY NONGROUP INS IN FEB15
66	PNGMA15	Num	4	COVERED BY NONGROUP INS IN MAR15
67	PNGAP15	Num	4	COVERED BY NONGROUP INS IN APR15
68	PNGMY15	Num	4	COVERED BY NONGROUP INS IN MAY15
69	PNGJU15	Num	4	COVERED BY NONGROUP INS IN JUN15
70	PNGJL15	Num	4	COVERED BY NONGROUP INS IN JUL15
71	PNGAU15	Num	4	COVERED BY NONGROUP INS IN AUG15
72	PNGSE15	Num	4	COVERED BY NONGROUP INS IN SEP15
73	PNGOC15	Num	4	COVERED BY NONGROUP INS IN OCT15
74	PNGNO15	Num	4	COVERED BY NONGROUP INS IN NOV15
75	PNGDE15	Num	4	COVERED BY NONGROUP INS IN DEC15
76	POGJA15	Num	4	COVERED BY OTHER GROUP INS IN JAN15
77	POGFE15	Num	4	COVERED BY OTHER GROUP INS IN FEB15
78	POGMA15	Num	4	COVERED BY OTHER GROUP INS IN MAR15
79	POGAP15	Num	4	COVERED BY OTHER GROUP INS IN APR15
80	POGMY15	Num	4	COVERED BY OTHER GROUP INS IN MAY15
81	POGJU15	Num	4	COVERED BY OTHER GROUP INS IN JUN15
82	POGJL15	Num	4	COVERED BY OTHER GROUP INS IN JUL15
83	POGAU15	Num	4	COVERED BY OTHER GROUP INS IN AUG15
84	POGSE15	Num	4	COVERED BY OTHER GROUP INS IN SEP15
85	POGOC15	Num	4	COVERED BY OTHER GROUP INS IN OCT15
86	POGNO15	Num	4	COVERED BY OTHER GROUP INS IN NOV15
87	POGDE15	Num	4	COVERED BY OTHER GROUP INS IN DEC15
88	PRSJA15	Num	4	COVERED BY SELF-EMP-1 INS IN JAN15
89	PRSFE15	Num	4	COVERED BY SELF-EMP-1 INS IN FEB15
90	PRSMA15	Num	4	COVERED BY SELF-EMP-1 INS IN MAR15
91	PRSAP15	Num	4	COVERED BY SELF-EMP-1 INS IN APR15
92	PRSMY15	Num	4	COVERED BY SELF-EMP-1 INS IN MAY15
93	PRSJU15	Num	4	COVERED BY SELF-EMP-1 INS IN JUN15
94	PRSJL15	Num	4	COVERED BY SELF-EMP-1 INS IN JUL15
95	PRSAU15	Num	4	COVERED BY SELF-EMP-1 INS IN AUG15
96	PRSSE15	Num	4	COVERED BY SELF-EMP-1 INS IN SEP15
97	PRSOC15	Num	4	COVERED BY SELF-EMP-1 INS IN OCT15
98	PRSNO15	Num	4	COVERED BY SELF-EMP-1 INS IN NOV15
99	PRSDE15	Num	4	COVERED BY SELF-EMP-1 INS IN DEC15
100	POUJA15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN JAN15
101	POUFE15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN FEB15
102	POUMA15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN MAR15
103	POUAP15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN APR15
104	POUMY15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN MAY15
105	POUJU15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN JUN15
106	POUJL15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN JUL15
107	POUAU15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN AUG15
108	POUSE15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN SEP15
109	POUOC15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN OCT15
110	POUNO15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN NOV15
111	POUDE15	Num	4	COVERED BY HOLDER OUTSIDE OF RU IN DEC15
112	PRXJA15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN JAN15
113	PRXFE15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN FEB15
114	PRXMA15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN MAR15
115	PRXAP15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN APR15
116	PRXMY15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN MAY15
117	PRXJU15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN JUN15
118	PRXJL15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN JUL15

119	PRXAU15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN AUG15
120	PRXSE15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN SEP15
121	PRXOC15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN OCT15
122	PRXNO15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN NOV15
123	PRXDE15	Num	4	COV BY PRIV INS THROUGH EXCHNG IN DEC15
124	PRIJA15	Num	4	COVERED BY PRIVATE INS IN JAN15
125	PRIFE15	Num	4	COVERED BY PRIVATE INS IN FEB15
126	PRIMA15	Num	4	COVERED BY PRIVATE INS IN MAR15
127	PRIAP15	Num	4	COVERED BY PRIVATE INS IN APR15
128	PRIMY15	Num	4	COVERED BY PRIVATE INS IN MAY15
129	PRIJU15	Num	4	COVERED BY PRIVATE INS IN JUN15
130	PRIJL15	Num	4	COVERED BY PRIVATE INS IN JUL15
131	PRIAU15	Num	4	COVERED BY PRIVATE INS IN AUG15
132	PRISE15	Num	4	COVERED BY PRIVATE INS IN SEP15
133	PRIOC15	Num	4	COVERED BY PRIVATE INS IN OCT15
134	PRINO15	Num	4	COVERED BY PRIVATE INS IN NOV15
135	PRIDE15	Num	4	COVERED BY PRIVATE INS IN DEC15
136	INSJA15X	Num	4	COVR BY HOSP/MED INS IN JAN15 (ED)
137	INSFE15X	Num	4	COVR BY HOSP/MED INS IN FEB15 (ED)
138	INSMA15X	Num	4	COVR BY HOSP/MED INS IN MAR15 (ED)
139	INSAP15X	Num	4	COVR BY HOSP/MED INS IN APR15 (ED)
140	INSMY15X	Num	4	COVR BY HOSP/MED INS IN MAY15 (ED)
141	INSJU15X	Num	4	COVR BY HOSP/MED INS IN JUN15 (ED)
142	INSJL15X	Num	4	COVR BY HOSP/MED INS IN JUL15 (ED)
143	INSAU15X	Num	4	COVR BY HOSP/MED INS IN AUG15 (ED)
144	INSSE15X	Num	4	COVR BY HOSP/MED INS IN SEP15 (ED)
145	INSOC15X	Num	4	COVR BY HOSP/MED INS IN OCT15 (ED)
146	INSNO15X	Num	4	COVR BY HOSP/MED INS IN NOV15 (ED)
147	INSDE15X	Num	4	COVR BY HOSP/MED INS IN DEC15 (ED)

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE7.SAS: HOW TO CONSTRUCT INSURANCE STATUS VARIABLES, USING FY 15 DATA
CREATE COUNT VARIABLES FOR # OF MONTHS WITH INSURANCE

The FREQ Procedure

OF MONTHS COV BY PRIVATE INSU

PRI_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	17182	48.50	17182	48.50
1	146	0.41	17328	48.91
2	169	0.48	17497	49.39
3	239	0.67	17736	50.06
4	306	0.86	18042	50.93
5	383	1.08	18425	52.01
6	313	0.88	18738	52.89
7	275	0.78	19013	53.67
8	271	0.76	19284	54.43
9	333	0.94	19617	55.37
10	431	1.22	20048	56.59
11	449	1.27	20497	57.86
12	14930	42.14	35427	100.00

OF MONTHS COV BY ANY INSU

INS_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	4234	11.95	4234	11.95
1	170	0.48	4404	12.43
2	228	0.64	4632	13.07
3	305	0.86	4937	13.94
4	375	1.06	5312	14.99
5	473	1.34	5785	16.33
6	374	1.06	6159	17.39
7	492	1.39	6651	18.77
8	466	1.32	7117	20.09
9	565	1.59	7682	21.68
10	688	1.94	8370	23.63
11	747	2.11	9117	25.73
12	26310	74.27	35427	100.00

OF MONTHS WITHOUT INSU

UNINS_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	27271	76.98	27271	76.98
1	724	2.04	27995	79.02
2	672	1.90	28667	80.92
3	534	1.51	29201	82.43
4	427	1.21	29628	83.63
5	403	1.14	30031	84.77
6	326	0.92	30357	85.69
7	355	1.00	30712	86.69
8	283	0.80	30995	87.49
9	239	0.67	31234	88.16
10	177	0.50	31411	88.66
11	123	0.35	31534	89.01
12	3893	10.99	35427	100.00

OF MONTHS COV BY MEDICAID

MCD_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	23821	67.24	23821	67.24
1	132	0.37	23953	67.61
2	194	0.55	24147	68.16
3	246	0.69	24393	68.85
4	238	0.67	24631	69.53
5	308	0.87	24939	70.40
6	257	0.73	25196	71.12
7	349	0.99	25545	72.11
8	375	1.06	25920	73.16
9	357	1.01	26277	74.17
10	419	1.18	26696	75.35
11	400	1.13	27096	76.48
12	8331	23.52	35427	100.00

OF MONTHS COV BY MEDICARE

MCR_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	30093	84.94	30093	84.94
1	24	0.07	30117	85.01
2	33	0.09	30150	85.10
3	37	0.10	30187	85.21
4	60	0.17	30247	85.38
5	60	0.17	30307	85.55
6	53	0.15	30360	85.70
7	38	0.11	30398	85.80
8	43	0.12	30441	85.93
9	29	0.08	30470	86.01
10	42	0.12	30512	86.13
11	57	0.16	30569	86.29
12	4858	13.71	35427	100.00

OF MONTHS COV BY TRICARE

TRI_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	34555	97.54	34555	97.54
1	13	0.04	34568	97.58
2	17	0.05	34585	97.62
3	21	0.06	34606	97.68
4	20	0.06	34626	97.74
5	18	0.05	34644	97.79
6	17	0.05	34661	97.84
7	27	0.08	34688	97.91
8	9	0.03	34697	97.94

9	14	0.04	34711	97.98
10	30	0.08	34741	98.06
11	30	0.08	34771	98.15
12	656	1.85	35427	100.00

OF MONTHS COV BY OTHER PUBLIC A OR B INSU

OPAB_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	34902	98.52	34902	98.52
1	26	0.07	34928	98.59
2	40	0.11	34968	98.70
3	37	0.10	35005	98.81
4	57	0.16	35062	98.97
5	55	0.16	35117	99.12
6	27	0.08	35144	99.20
7	27	0.08	35171	99.28
8	24	0.07	35195	99.35
9	18	0.05	35213	99.40
10	33	0.09	35246	99.49
11	31	0.09	35277	99.58
12	150	0.42	35427	100.00

OF MONTHS COV BY PRIVATE GROUP INSU

GRP_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	18705	52.80	18705	52.80
1	135	0.38	18840	53.18
2	178	0.50	19018	53.68
3	214	0.60	19232	54.29
4	261	0.74	19493	55.02
5	355	1.00	19848	56.03
6	285	0.80	20133	56.83
7	280	0.79	20413	57.62
8	249	0.70	20662	58.32
9	296	0.84	20958	59.16
10	329	0.93	21287	60.09
11	347	0.98	21634	61.07
12	13793	38.93	35427	100.00

OF MONTHS COV BY PRIVATE NON-GROUP INSU

NG_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	32915	92.91	32915	92.91
1	63	0.18	32978	93.09
2	51	0.14	33029	93.23
3	83	0.23	33112	93.47
4	132	0.37	33244	93.84
5	137	0.39	33381	94.22
6	95	0.27	33476	94.49
7	63	0.18	33539	94.67
8	84	0.24	33623	94.91
9	78	0.22	33701	95.13
10	138	0.39	33839	95.52
11	131	0.37	33970	95.89
12	1457	4.11	35427	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE7.SAS: HOW TO CONSTRUCT INSURANCE STATUS VARIABLES, USING FY 15 DATA
CREATE COUNT VARIABLES FOR # OF MONTHS WITH INSURANCE

The FREQ Procedure

OF MONTHS COV BY PUBLIC INSU

PUB_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	19548	55.18	19548	55.18
1	151	0.43	19699	55.60
2	212	0.60	19911	56.20
3	259	0.73	20170	56.93
4	274	0.77	20444	57.71
5	334	0.94	20778	58.65
6	264	0.75	21042	59.40
7	356	1.00	21398	60.40
8	376	1.06	21774	61.46
9	353	1.00	22127	62.46
10	430	1.21	22557	63.67
11	428	1.21	22985	64.88
12	12442	35.12	35427	100.00

OF MONTHS IN MEPS SURVEY

REF_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	88	0.25	88	0.25
2	117	0.33	205	0.58
3	136	0.38	341	0.96
4	130	0.37	471	1.33
5	148	0.42	619	1.75
6	149	0.42	768	2.17
7	142	0.40	910	2.57
8	124	0.35	1034	2.92
9	137	0.39	1171	3.31
10	123	0.35	1294	3.65
11	117	0.33	1411	3.98
12	34016	96.02	35427	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE7.SAS: HOW TO CONSTRUCT INSURANCE STATUS VARIABLES, USING FY 15 DATA
SAMPLE DUMP TO CHECK THE COUNT VARIABLES (USING PRI_N AS EXMAPLE)

DUPERSID	PRIJA15	PRIFE15	PRIMA15	PRIAP15	PRIMY15	PRIJU15	PRIJL15	PRIAU15	PRISE15	PRIOC15	PRINO15	PRIDE15	PRI_N
60001101	2	2	2	2	2	2	2	2	2	2	2	2	0
60001102	2	2	2	2	2	2	2	2	2	2	2	2	0
60001103	1	1	1	1	1	1	1	1	1	1	1	1	12
60001104	2	2	2	2	2	2	2	2	2	2	2	2	0
60001105	2	2	2	2	2	2	2	2	2	2	2	2	0
60002101	2	2	2	2	2	2	2	2	2	2	2	2	0
60002102	2	2	2	2	2	2	2	2	2	2	2	2	0
60002103	2	2	2	2	2	2	2	2	2	2	2	2	0
60002104	2	2	2	2	2	2	2	2	2	2	2	2	0
60003101	1	1	1	1	1	1	1	1	1	1	1	1	12
60003102	1	1	1	1	1	1	1	1	1	1	1	1	12
60003103	1	1	1	1	1	1	1	1	1	1	1	1	12
60003104	1	1	1	1	1	1	1	1	1	1	1	1	12
60004101	1	1	1	1	1	1	1	1	1	1	1	1	12
60004102	2	2	2	2	2	2	2	2	2	2	2	2	0
60004103	2	2	2	2	2	2	2	2	2	2	2	2	0
60009101	1	1	1	1	1	1	1	1	1	1	1	1	12
60010101	2	2	1	1	1	1	1	1	1	1	1	1	10
60010103	2	2	2	2	2	2	2	2	2	2	2	2	0
60010106	2	2	2	2	2	2	2	2	2	2	2	2	0
60010107	2	2	2	2	2	2	2	2	2	2	2	2	0
60011101	2	1	1	1	1	1	1	1	1	1	1	1	11
60011102	2	1	1	1	1	1	1	1	1	1	1	1	11
60011103	2	1	1	1	1	1	1	1	1	1	1	1	11
60011104	2	1	1	1	1	1	1	1	1	1	1	1	11
60013101	1	1	1	1	1	1	1	1	1	1	1	1	12
60013102	1	1	1	1	1	1	1	1	1	1	1	1	12
60013103	1	1	1	1	1	1	1	1	1	1	1	1	12
60015101	2	2	2	2	2	2	2	2	2	1	1	1	3
60015102	1	1	1	1	1	1	1	1	1	1	1	1	12

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE7.SAS: HOW TO CONSTRUCT INSURANCE STATUS VARIABLES, USING FY 15 DATA
SUPPORTING CROSSTABS TO VERIFY THE CREATION OF THE FLAGS

The FREQ Procedure

INSURED FOR FULL YEAR

FULL_INSU	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	8156	23.02	8156	23.02
1	27271	76.98	35427	100.00

FULL_INSU	UNINS_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1	724	2.04	724	2.04
0	2	672	1.90	1396	3.94
0	3	534	1.51	1930	5.45
0	4	427	1.21	2357	6.65
0	5	403	1.14	2760	7.79
0	6	326	0.92	3086	8.71
0	7	355	1.00	3441	9.71
0	8	283	0.80	3724	10.51
0	9	239	0.67	3963	11.19
0	10	177	0.50	4140	11.69
0	11	123	0.35	4263	12.03
0	12	3893	10.99	8156	23.02
1	0	27271	76.98	35427	100.00

FULL_INSU	INS_N	REF_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	0	1	31	0.09	31	0.09
0	0	2	45	0.13	76	0.21
0	0	3	47	0.13	123	0.35
0	0	4	38	0.11	161	0.45
0	0	5	24	0.07	185	0.52
0	0	6	48	0.14	233	0.66
0	0	7	16	0.05	249	0.70
0	0	8	15	0.04	264	0.75
0	0	9	31	0.09	295	0.83
0	0	10	28	0.08	323	0.91
0	0	11	18	0.05	341	0.96
0	0	12	3893	10.99	4234	11.95
0	1	2	3	0.01	4237	11.96
0	1	4	2	0.01	4239	11.97
0	1	7	1	0.00	4240	11.97
0	1	8	1	0.00	4241	11.97
0	1	10	1	0.00	4242	11.97
0	1	12	105	0.30	4347	12.27
0	2	3	1	0.00	4348	12.27
0	2	4	3	0.01	4351	12.28
0	2	5	1	0.00	4352	12.28
0	2	7	1	0.00	4353	12.29
0	2	8	1	0.00	4354	12.29
0	2	10	1	0.00	4355	12.29
0	2	11	2	0.01	4357	12.30
0	2	12	149	0.42	4506	12.72
0	3	4	1	0.00	4507	12.72
0	3	6	1	0.00	4508	12.72
0	3	7	3	0.01	4511	12.73
0	3	9	6	0.02	4517	12.75
0	3	10	1	0.00	4518	12.75
0	3	12	205	0.58	4723	13.33
0	4	5	5	0.01	4728	13.35
0	4	6	5	0.01	4733	13.36
0	4	7	3	0.01	4736	13.37
0	4	8	4	0.01	4740	13.38
0	4	9	1	0.00	4741	13.38

0	4	11	4	0.01	4745	13.39
0	4	12	267	0.75	5012	14.15
0	5	6	3	0.01	5015	14.16
0	5	7	4	0.01	5019	14.17
0	5	8	3	0.01	5022	14.18
0	5	9	7	0.02	5029	14.20
0	5	10	2	0.01	5031	14.20
0	5	11	3	0.01	5034	14.21
0	5	12	333	0.94	5367	15.15
0	6	7	5	0.01	5372	15.16
0	6	8	3	0.01	5375	15.17
0	6	10	4	0.01	5379	15.18
0	6	11	3	0.01	5382	15.19
0	6	12	267	0.75	5649	15.95
0	7	8	6	0.02	5655	15.96
0	7	9	4	0.01	5659	15.97
0	7	10	1	0.00	5660	15.98
0	7	12	372	1.05	6032	17.03
0	8	9	1	0.00	6033	17.03
0	8	10	2	0.01	6035	17.04
0	8	11	1	0.00	6036	17.04
0	8	12	371	1.05	6407	18.09
0	9	10	3	0.01	6410	18.09
0	9	12	475	1.34	6885	19.43
0	10	11	2	0.01	6887	19.44
0	10	12	606	1.71	7493	21.15
0	11	12	663	1.87	8156	23.02
1	1	1	57	0.16	8213	23.18
1	2	2	69	0.19	8282	23.38
1	3	3	88	0.25	8370	23.63
1	4	4	86	0.24	8456	23.87
1	5	5	118	0.33	8574	24.20
1	6	6	92	0.26	8666	24.46
1	7	7	109	0.31	8775	24.77
1	8	8	91	0.26	8866	25.03
1	9	9	87	0.25	8953	25.27
1	10	10	80	0.23	9033	25.50
1	11	11	84	0.24	9117	25.73
1	12	12	26310	74.27	35427	100.00

EVER INSURED BY PRIVATE GROUP

GROUP_INS1	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	18705	52.80	18705	52.80
1	16722	47.20	35427	100.00

GROUP_INS1	GRP_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	0	18705	52.80	18705	52.80
1	1	135	0.38	18840	53.18
1	2	178	0.50	19018	53.68
1	3	214	0.60	19232	54.29
1	4	261	0.74	19493	55.02
1	5	355	1.00	19848	56.03
1	6	285	0.80	20133	56.83
1	7	280	0.79	20413	57.62
1	8	249	0.70	20662	58.32
1	9	296	0.84	20958	59.16
1	10	329	0.93	21287	60.09
1	11	347	0.98	21634	61.07
1	12	13793	38.93	35427	100.00

INSURED BY PRIVATE GROUP FOR FULL YEAR

GROUP_INS2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	21272	60.04	21272	60.04
1	14155	39.96	35427	100.00

GROUP_INS2	GRP_N	REF_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	0	1	69	0.19	69	0.19
0	0	2	88	0.25	157	0.44
0	0	3	101	0.29	258	0.73
0	0	4	95	0.27	353	1.00
0	0	5	97	0.27	450	1.27
0	0	6	93	0.26	543	1.53
0	0	7	93	0.26	636	1.80
0	0	8	86	0.24	722	2.04
0	0	9	88	0.25	810	2.29
0	0	10	88	0.25	898	2.53
0	0	11	85	0.24	983	2.77
0	0	12	17722	50.02	18705	52.80
0	1	2	2	0.01	18707	52.80
0	1	4	1	0.00	18708	52.81
0	1	5	4	0.01	18712	52.82
0	1	9	1	0.00	18713	52.82
0	1	12	108	0.30	18821	53.13
0	2	3	1	0.00	18822	53.13
0	2	4	1	0.00	18823	53.13
0	2	5	1	0.00	18824	53.13
0	2	8	1	0.00	18825	53.14
0	2	10	2	0.01	18827	53.14
0	2	12	145	0.41	18972	53.55
0	3	4	1	0.00	18973	53.56
0	3	5	1	0.00	18974	53.56
0	3	7	2	0.01	18976	53.56
0	3	9	1	0.00	18977	53.57
0	3	10	1	0.00	18978	53.57
0	3	12	174	0.49	19152	54.06
0	4	5	3	0.01	19155	54.07
0	4	6	2	0.01	19157	54.07
0	4	7	3	0.01	19160	54.08
0	4	8	2	0.01	19162	54.09
0	4	10	1	0.00	19163	54.09
0	4	11	4	0.01	19167	54.10
0	4	12	214	0.60	19381	54.71
0	5	6	2	0.01	19383	54.71
0	5	7	4	0.01	19387	54.72
0	5	8	1	0.00	19388	54.73
0	5	9	6	0.02	19394	54.74
0	5	10	3	0.01	19397	54.75
0	5	11	2	0.01	19399	54.76
0	5	12	295	0.83	19694	55.59
0	6	7	1	0.00	19695	55.59
0	6	8	1	0.00	19696	55.60
0	6	9	1	0.00	19697	55.60
0	6	10	1	0.00	19698	55.60

GROUP_INS2	GRP_N	REF_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	6	12	229	0.65	19927	56.25
0	7	8	4	0.01	19931	56.26
0	7	9	2	0.01	19933	56.26
0	7	12	235	0.66	20168	56.93
0	8	9	1	0.00	20169	56.93
0	8	10	1	0.00	20170	56.93
0	8	12	218	0.62	20388	57.55
0	9	11	1	0.00	20389	57.55
0	9	12	258	0.73	20647	58.28
0	10	12	303	0.86	20950	59.14
0	11	12	322	0.91	21272	60.04
1	1	1	19	0.05	21291	60.10
1	2	2	27	0.08	21318	60.17

1	3	3	34	0.10	21352	60.27
1	4	4	32	0.09	21384	60.36
1	5	5	42	0.12	21426	60.48
1	6	6	52	0.15	21478	60.63
1	7	7	39	0.11	21517	60.74
1	8	8	29	0.08	21546	60.82
1	9	9	37	0.10	21583	60.92
1	10	10	26	0.07	21609	61.00
1	11	11	25	0.07	21634	61.07
1	12	12	13793	38.93	35427	100.00

EVER INSURED BY PRIVATE NON-GROUP

NG_INS	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	32915	92.91	32915	92.91
1	2512	7.09	35427	100.00

NG_INS	NG_N	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	0	32915	92.91	32915	92.91
1	1	63	0.18	32978	93.09
1	2	51	0.14	33029	93.23
1	3	83	0.23	33112	93.47
1	4	132	0.37	33244	93.84
1	5	137	0.39	33381	94.22
1	6	95	0.27	33476	94.49
1	7	63	0.18	33539	94.67
1	8	84	0.24	33623	94.91
1	9	78	0.22	33701	95.13
1	10	138	0.39	33839	95.52
1	11	131	0.37	33970	95.89
1	12	1457	4.11	35427	100.00

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE7.SAS: HOW TO CONSTRUCT INSURANCE STATUS VARIABLES, USING FY 15 DATA
% AND POPULATION WITH INSU

The SURVEYMEANS Procedure

Data Summary

Number of Strata	165
Number of Clusters	369
Number of Observations	35427
Number of Observations Used	33893
Number of Obs with Nonpositive Weights	1534
Sum of Weights	321423251

Statistics

Variable	Label	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
FULL_INSU	INSURED FOR FULL YEAR	33893	321423251	0.822103	0.004876	264243123	6380117
GROUP_INS1	EVER INSURED BY PRIVATE GROUP	33893	321423251	0.594406	0.008005	191056055	5776071
GROUP_INS2	INSURED BY PRIVATE GROUP FOR FULL YEAR	33893	321423251	0.516621	0.007646	166053963	5127080
NG_INS	EVER INSURED BY PRIVATE NON-GROUP	33893	321423251	0.090758	0.003628	29171728	1292574

2018 AHRQ MEPS DATA USERS WORKSHOP
EXERCISE7.SAS: HOW TO CONSTRUCT INSURANCE STATUS VARIABLES, USING FY 15 DATA
% AND POPULATION WITH INSU

The SURVEYMEANS Procedure

Statistics for RACETHX Domains

RACETHX	Variable	Label	N	Sum of Weights	Mean	Std Error of Mean	Sum	Std Error of Sum
1 HISPANIC	FULL_INSU	INSURED FOR FULL YEAR	11068	57063251	0.682578	0.010694	38950120	2229921
	GROUP_INS1	EVER INSURED BY PRIVATE GROUP	11068	57063251	0.419398	0.014257	23932241	1499004

	GROUP_INS2	INSURED BY PRIVATE GROUP FOR FULL YEAR	11068	57063251	0.347345	0.012612	19820642	1248024
	NG_INS	EVER INSURED BY PRIVATE NON-GROUP	11068	57063251	0.046007	0.003745	2625282	275959
2 WHITE	FULL_INSU	INSURED FOR FULL YEAR	12565	194013304	0.867428	0.005643	168292666	5313767
	GROUP_INS1	EVER INSURED BY PRIVATE GROUP	12565	194013304	0.662127	0.009217	128461479	4868012
	GROUP_INS2	INSURED BY PRIVATE GROUP FOR FULL YEAR	12565	194013304	0.584405	0.008925	113382282	4386287
	NG_INS	EVER INSURED BY PRIVATE NON-GROUP	12565	194013304	0.112172	0.005628	21762770	1194577
3 BLACK	FULL_INSU	INSURED FOR FULL YEAR	6551	39532988	0.784068	0.007889	30996541	1642401
	GROUP_INS1	EVER INSURED BY PRIVATE GROUP	6551	39532988	0.490022	0.014922	19372035	1048198
	GROUP_INS2	INSURED BY PRIVATE GROUP FOR FULL YEAR	6551	39532988	0.402067	0.012814	15894924	891708
	NG_INS	EVER INSURED BY PRIVATE NON-GROUP	6551	39532988	0.053230	0.003674	2104322	162818
4 ASIAN	FULL_INSU	INSURED FOR FULL YEAR	2464	17835787	0.855344	0.011147	15255729	1412252
	GROUP_INS1	EVER INSURED BY PRIVATE GROUP	2464	17835787	0.635319	0.023777	11331412	1167388
	GROUP_INS2	INSURED BY PRIVATE GROUP FOR FULL YEAR	2464	17835787	0.573825	0.023911	10234624	1088710
	NG_INS	EVER INSURED BY PRIVATE NON-GROUP	2464	17835787	0.118399	0.014517	2111742	264823
5 OTHER RACE	FULL_INSU	INSURED FOR FULL YEAR	1245	12977919	0.828181	0.023337	10748068	1069816
	GROUP_INS1	EVER INSURED BY PRIVATE GROUP	1245	12977919	0.613264	0.026607	7958888	840897
	GROUP_INS2	INSURED BY PRIVATE GROUP FOR FULL YEAR	1245	12977919	0.517917	0.038913	6721491	665306
	NG_INS	EVER INSURED BY PRIVATE NON-GROUP	1245	12977919	0.043737	0.007879	567613	103025

Output End

MEPS Workshop Report

As can be seen from the above scripts and output, The MEPS Workshop provides in-depth analytical techniques and use cases for the MEPS datasets. Let us look at what is going on in each of the exercises.

Exercise 1.A

In analyzing the results of Exercise 1.A we see from the initial descriptive statistics that we are analyzing total expenditures by age (0-64, 65+) for the year 2016. Our dataset includes 34,655 responses. In the cross-tabulation table for the flag variables, we see 6,933 people did not have a healthcare related expenditure in 2016 (20.01%), while 27,722 did report a healthcare related expenditure (79.99%). Likewise, 29,910 people interviewed were between the ages of 0-64 (86.31%), while 4,745 of those interviewed were 65+ (13.69%).

We first use the PROC SURVEYMEANS statement to estimate the number of noninstitutionalized U.S. civilians with a healthcare expenditure in 2016. We see this to be 273,773,024 individuals, or approximately 85% of the entire population at 323,141,637 individuals. This analysis is based off 33,259 of our initial 34,655 individuals. Any observations with missing values for design variables such as STRATA, CLUSTER, or DOMAIN, will be excluded from the analysis if the MISSING option is not specified ("PROC SURVEYMEANS: Missing Values: SAS/STAT(R) 9.2 User's Guide, Second Edition," n.d.). Additionally, we estimate the error of our estimation to be 6,396,258 individuals.

Second, we perform the same analysis, though this time on the TOTAL variable as opposed to the X_ANYSVCE variable. This analysis allows us to estimate a mean healthcare expense of $5,005.64 for a noninstitutionalized U.S. citizen in 2016, give or take $116.95 (the

standard error of the mean). We estimate the total cost of healthcare expenditures for the population to be $1,617,531,007,314, or approximately $1.6 trillion, give or take $50,638,644,399, or approximately $50 billion.

Next, we provide a similar analysis, this time using domains. We are analyzing the domains of all persons with an expenditure (26,942 out of 27,772 possible responses), as well as all persons with an expenditure by age group (0 – 64, 65+). Out of 26,942 persons in our survey sample with a healthcare related expenditure, 22,513 are between the ages 0 - 64, while 4,429 are 65 or older.

We estimate the average expenditure for the sub-population of all U.S. citizens *with* an expenditure (273,773,024 persons) to be $5,908.30 per person. We estimate the average expenditure of all U.S. citizens ages 0 – 64 with an expenditure (224,366,322 persons) to be $4,615.10 per person. Finally, we estimate the average expenditure of all U.S. citizens ages 65 or older with an expenditure (49,406,702 persons) to be $11,780.80 per person.

There seems to be a significant gap ($7,165) between the amount persons ages 0-64 spend on healthcare and the amount persons 65 or older spend on healthcare. We use the PROC TTEST statement to perform a two-sample t-test on these given sub-populations, analyzing the expenditure variable (Elliott & Woodward, 2016). By providing the appropriate WEIGHT and WHERE clauses, we can pinpoint the correct sub-population, which we verify by confirming the number of observations (nobs=22,513 and nobs=4,429) and mean expenditures ($4,615.10 and $11,780.80) equal those for our designated sub-populations. Additionally, we confirm the difference between means is the same ($7,165).

The t-tests for both the pooled version and the Satterthwaite version produce a p-value < 0.0001. We can reject the null hypothesis that TOTAL expenditures between these two age groups (0-64, 65+) are equal; there is evidence for a relationship between age group (0-64, 65+) and estimated healthcare expenditures for the noninstitutionalized U.S. population. From the results of our F-test, we can conclude the variances between these sample populations are not equal, as the F-statistic produces a p-value < 0.0001. This is irrelevant, however, as the results of both of our two-sampled t-tests (pooled version and Satterthwaite version) produced t-statistics with p-values < 0.0001 (Elliott & Woodward, 2016).

Exercise 1.B

Exercise 1.B builds on the foundation set forth by Exercise 1.A, utilizing a similar pattern to analyze total expenditures by age combined with type of service. Type of service includes hospital inpatient services, ambulatory services, prescribed medication services, dental services, emergency room services, and home health care services. A quick glance at the supporting cross-tabulations for our flag variables tell us most respondents report no expense ($0) for hospital inpatient services, dental services, and home health services, while most respondents report an expense (>$0) for ambulatory services and prescribed medication services.

The rest of the results are identical to those produced above, though dependent upon each type of service. For instance, we estimate the average dental expense for any given person in the U.S. population to be $298. We estimate the average dental expense for any given person with a dental expense (134,304,354 persons) to be $714, give or take $17 (the standard error of the mean). When further analyzing these results by age group (0-64, 65+), we estimate the average dental expense of individuals aged 0-64 (111,592,123 persons) to be ~$670, while we estimate the average dental expense of individuals aged 65+ (22,712,232 persons) to be ~$930. Another

two-sample t-test can verify the significance of these results. Such a t-test produces a t-statistic with a p-value < 0.0001. Again, we can reject the null hypothesis and conclude there is evidence for a significant relationship between a person's age and expenditures for dental services (Elliott & Woodward, 2016).

Exercise 2.A

In Exercise 2.A we analyze expenditures as they relate to a specific type of prescribed medication: antipsychotic medications. Viewing the data at the level of the individual allows us to view the number of purchases, total cost of purchases, total out-of-pocket (OOP) expenditures, and total overall expenditures for third parties for any given individual. In examining the cross-tabulation analysis for our flag variable (persons with 1+ antipsychotic drugs), we see most respondents report both an OOP expense and a third-party expense, while slightly fewer report only third-party expenses.

When providing estimates for prescribed antipsychotic medication expenditures for the U.S. noninstitutionalized population, we estimate 4,464,275 persons as having an expense related to antipsychotic medications with the average total expense per person estimated to be $2,293.91. Likewise, we estimate an average of 6.9 purchases per person, with OOP expenses averaging $123.72 per person and third-party expenses averaging $2,170.48 per person.

Exercise 2.B

Exercise 2.B continues where the previous exercise left off, this time examining the estimated expenditures for narcotic analgesics, more commonly referred to as opiates. The results of our PROC SURVEYMEANS statement tells us an estimated 34,060,347 U.S. citizens are estimated to have an expenditure related to narcotic analgesics in 2016. We estimate the

average number of purchases per person for this domain to be 3.6 purchases per person, with an estimated average expense of $238.70 per person. We estimate $197.00 of this expense to be covered by third parties, while we estimate individuals to incur a $41.80 expenditure on average for purchases related to prescribed narcotic analgesics.

Exercise 3.A

In Exercise 3.A we analyze conditions and how they affect healthcare uses and expenditures. Specifically, we are analyzing the diabetic condition, and further analyzing this domain by sex. The unweighted results of the PROC FREQ statement reveal 3,065 persons in our survey sample with diabetes (13.69%), and 32,362 persons in our survey sample without the condition (91.35%). When breaking this analysis down further by sex, we estimate 13,531,921 males to have the condition (4.21%), and 143,640,000 males to not have the condition (44.69%). Similarly, we estimate 14,168,669 females to have the condition (4.41%), and 150,090,000 females to not have the condition (46.69%).

When estimating the annual healthcare uses and expenditures as they relate to the diabetic condition, we estimate males with the condition to make an average of 11 office visits per year, and females with the condition to make an average of 12.58 office visits per year. We estimate the average total cost to males with the condition to be $14,419.19 with an estimated OOP expenditure of $1,203.37 per person and the average total cost to females with the condition to be $13,968.05 with an estimated OOP expenditure of $1,189.04 per person.

Exercise 3.B

Exercise 3.B continues where the previous exercise left off, this time calculating the expenditures for all events associated with a condition. We continue to analyze diabetes in our

example. An event consists of being prescribed medication, making an office-based visit, an emergency room visit, an inpatient visit (when a person is actually admitted into a facility), an outpatient visit (when a person is sent to an offsite location), or a home health visit.

We reference data on each of these events via separate files and link them with our condition file via a condition-event linking file. The event files also contain information on the source of payment, including family, Medicare, Medicaid, private insurance, veterans, state & local government, or workers compensation, among others. After linking, merging, and de-duplicating our observations, we can see for a given event type (e.g., ambulatory, prescribed medication, etc.), the amount of payment associated with that event type by payment source. For example, we can see that Medicare paid $1,287.28 out of $1,290.40 for an ambulatory event associated with diabetes for an individual. The remaining $3.12 was covered by Medicaid.

Likewise, we can go as far as to estimate the expenditure for an individual based on a given condition (e.g., diabetes), a given event-type (e.g., prescribed medication, office-based visit, inpatient visit, outpatient visit, etc.) and a given source of payment (e.g., family, Medicare, Medicaid, private insurance, etc.). For instance, we estimate individuals with diabetes to make an average of 1.6 emergency room visits per year due to the condition. We estimate the total costs of these visits to be $1,868.18 per person, with an estimated OOP expenditure to be approximately $178.31 per person. We estimate private insurance to cover $775.31 of these expenses.

Exercise 4.A

In Exercise 4.A we analyze individuals based on income level. Specifically, we analyze a sub-population aged 26-30 who are uninsured but have high income. We start by using the

PROC FREQ statement to form a cross-tabulation table where we can see frequency counts of income level by insurance coverage status. For instance, we can see in our survey sample most individuals ranked as poor (poor is determined as income level as percent of the poverty line) ("MEPS HC-192 2016 Full Year Consolidated Data File," n.d.), are receiving 'public only' insurance coverage. The second most popular insurance coverage status for those determined as poor is 'uninsured'. Contrast this with the most popular insurance coverage status for individuals ranked as high income or middle income: 'any private insurance'. In our example, we are analyzing 26-30 year old persons who are uninsured but have high income. Using the PROC SURVEYMEANS statement we can estimate the total expenditures for this sub-population: $241.50.

Exercise 4.B

Exercise 4.B continues our analysis of this sub-population using longitudinal files, which are files presenting data over a span of multiple years. We can use the PROC FREQ statement to see our sub-population represents just a tiny sliver (0.15%) of our overall survey sample. Further output from the PROC FREQ statement allows us to analyze the flag variable for our sub-population in coordination with our categorical variables (poverty status, insurance status, and age group). Lastly, we use the PROC SURVEYMEANS statement combined with the PROC PRINT statement to project the insurance coverage status for our sub-population (Age: 26-30, uninsured whole year, high income) into the following year. For example, we estimate 0.23 percent of our sub-population to obtain private insurance status within the following year, .03 percent to obtain public only insurance status, and .73 percent to remain uninsured.

Exercise 5.A

In Exercise 5.A we learn to analyze family units, and how families are represented in the MEPS dataset. We learn "The CPSFAMID represents a redefinition of MEPS families into families defined by the Current Population Survey (CPS). Some of the distinctions between CPS- and MEPS-defined families are that MEPS families include, and CPS families do not include non-married partners, foster children, and in-laws. These persons are considered as members of separate families for CPS-like families" ("MEPS HC-181 2015 Full Year Consolidated Data File," n.d.). We make use of the PROC PRINT statement to observe, in the output, how families are represented when we de-duplicate our results and group them by the corresponding dwelling-unit ID (DUID) and CPSFAMID. Lastly, we can estimate the total number of CPS families in the noninstitutionalized U.S. population as being 141,808,362 families and the average number of people per CPS family being 2.20 persons. Similarly, we estimate total OOP expenses for CPS families to be $1,328.59 and we estimate total income for CPS families to be $64,690 per year.

Exercise 5.B

The last exercise in the MEPS workshop illustrates how to construct insurance status variables from monthly insurance variables in the person-level data. Our cross-tabulation tables display the total number of months covered for a given insurance coverage type (Tricare, Medicare, Medicaid, Private Group Insurance, Private Non-Group Insurance, Public Insurance, etc.). We see, for instance, that 17,182 persons (48.50%) in our survey sample, are covered by private insurance for 0 months, while 14,930 persons (42.14%) report being covered by private insurance for 12 months. Similarly, we see 30,093 individuals (84.94%) are covered by Medicare for 0 months, while 4,858 (13.71%) report being covered by Medicare for 12 months.

The full spectrum of month counts by insurance type are available to us for all the months for a corresponding year, ranging from 0 – 12 months.

Once again, we utilize the PROC PRINT statement to display a sample dump with insurance status month flags, representing whether an individual was covered by private insurance in any given month of the corresponding year. For example, a value of "2" in the field PRIJA15, tells us an individual was not covered by private insurance in the month of January of the year 2015, while a value of "1" in the field PRIDE15 tells us an individual was covered by private insurance for the month of December, in the year 2015. Monthly insurance variables exist for each insurance coverage type (e.g., PNGMA15 represents whether an individual was covered by nongroup insurance in March 2015, while PRSOC15 represents whether an individual was covered by self-employed insurance in October 2015) ("Medical Expenditure Panel Survey PUF Codebook," n.d.).

Using this knowledge, we create 4 sub-population flags to determine an individual's insurance coverage status: 'insured for full year', 'ever insured by private group', 'insured by private group for full year', and 'ever insured by private non-group.' We can, for instance, determine if an individual was insured for the full year, by summing all the uninsured months for a given individual. If an individual is recorded as being without any insurance for any given month (represented by a "2" in one of the corresponding INS pre-fixed variables), we add a count of "1" to a counter representing number of months uninsured for any given individual. If an individual's count of uninsured months = 0, we can say he or she was insured for the full year.

We determine whether an individual was insured by private group insurance for the full year by calculating whether the number of months the participant has been in the survey is equal to the number of months the participant has been covered by private group insurance. Private

group insurance coverage is an aggregate representation composed of Tricare insurance coverage (TRI—15), outside holder coverage (POU—15), employer union coverage (PEG—15), and private coverage with source unknown (PDK—15) ("Medical Expenditure Panel Survey PUF Codebook," n.d.).

The last step in this exercise allows us to estimate the percent of a sub-population that we expect to see in a given coverage status for a corresponding year. For instance, we can estimate the percent of individuals we expect to see in each coverage status by race. A quick glance at our results tells us that we expect 68% of Hispanics to be fully insured for the year, 86% of Caucasians to be fully insured for the year, 78% of African-Americans to be fully insured for the year, 85% of Asians to be fully insured for the year, and 82% of all other races to be fully insured for the year.

Conclusion

In order to complete this project, the author utilized the directions found on the GitHub website to clone the corresponding repository and update the script files with explanatory comments as needed ("HHS-AHRQ/MEPS," n.d.). The author made extensive use of the SAS documentation as well as the MEPS website to research any unfamiliar terms and concepts in both the SAS code and the MEPS datafiles. The author used GitHub to provide source control for his modified code files and considered several different ways to convey the SAS output in a Microsoft Word document, before ultimately deciding to use text output from the log file. The author produced extensive documentation for each section of SAS code for all corresponding workshop exercises. Additionally, the author extended the workshop exercises by providing code to produce descriptive statistics and two-sample t-tests. The following additional screenshots provide a further showcase of the author's analysis of the MEPS datasets.

Additional Screenshots

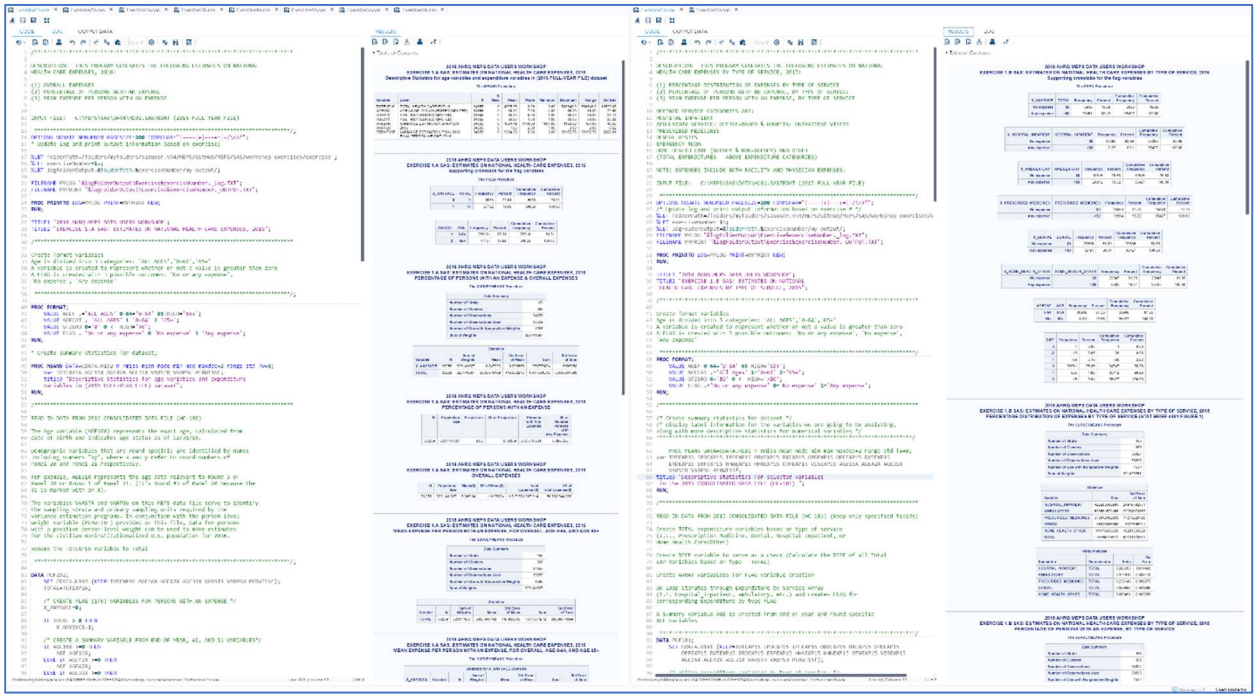

Figure 3. Exercise 1.A and Exercise 1.B script and output

Figure 4. Exercise 2.A and Exercise 2.B output

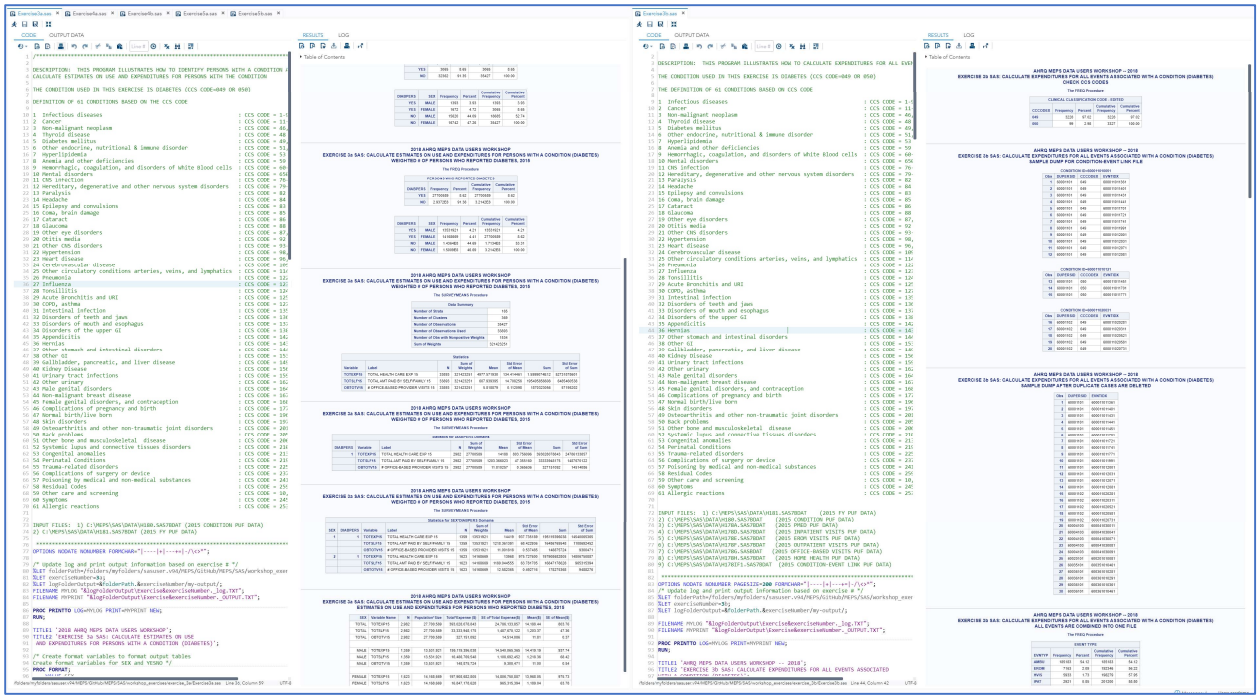

Figure 5. Exercise 3.A and Exercise 3.B output

Figure 6. Exercise 4.A and Exercise 4.B output

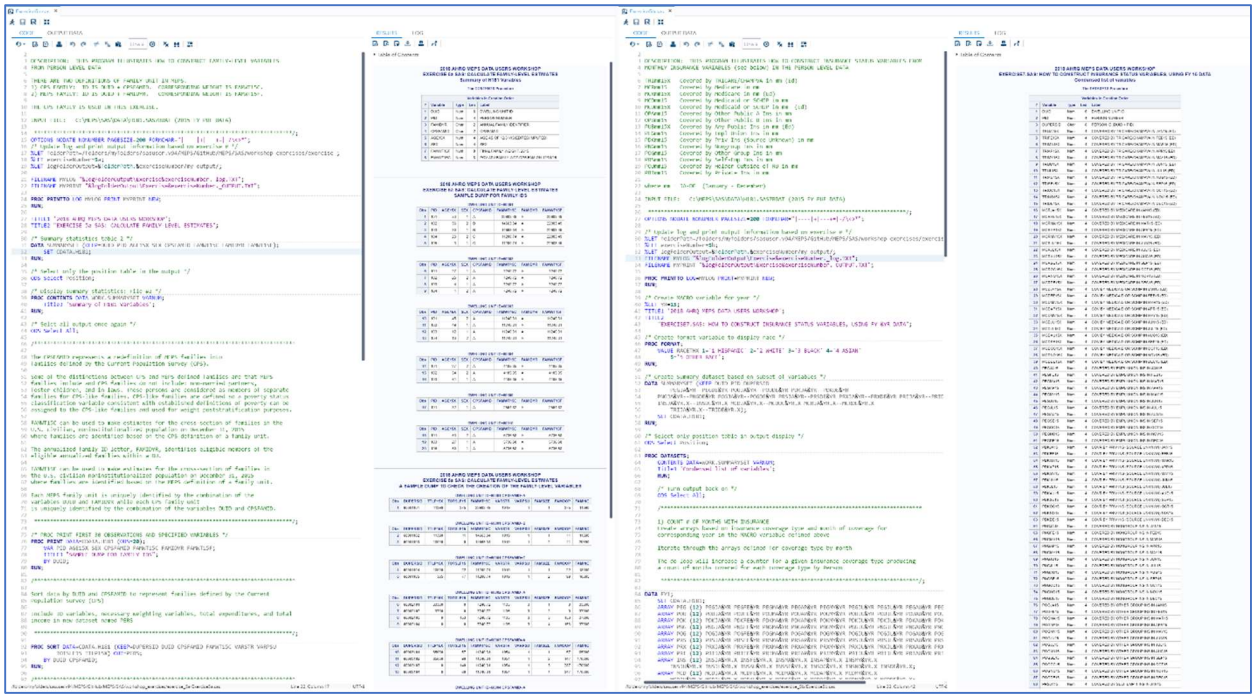

Figure 7. Exercise 5.A and Exercise 5.B output

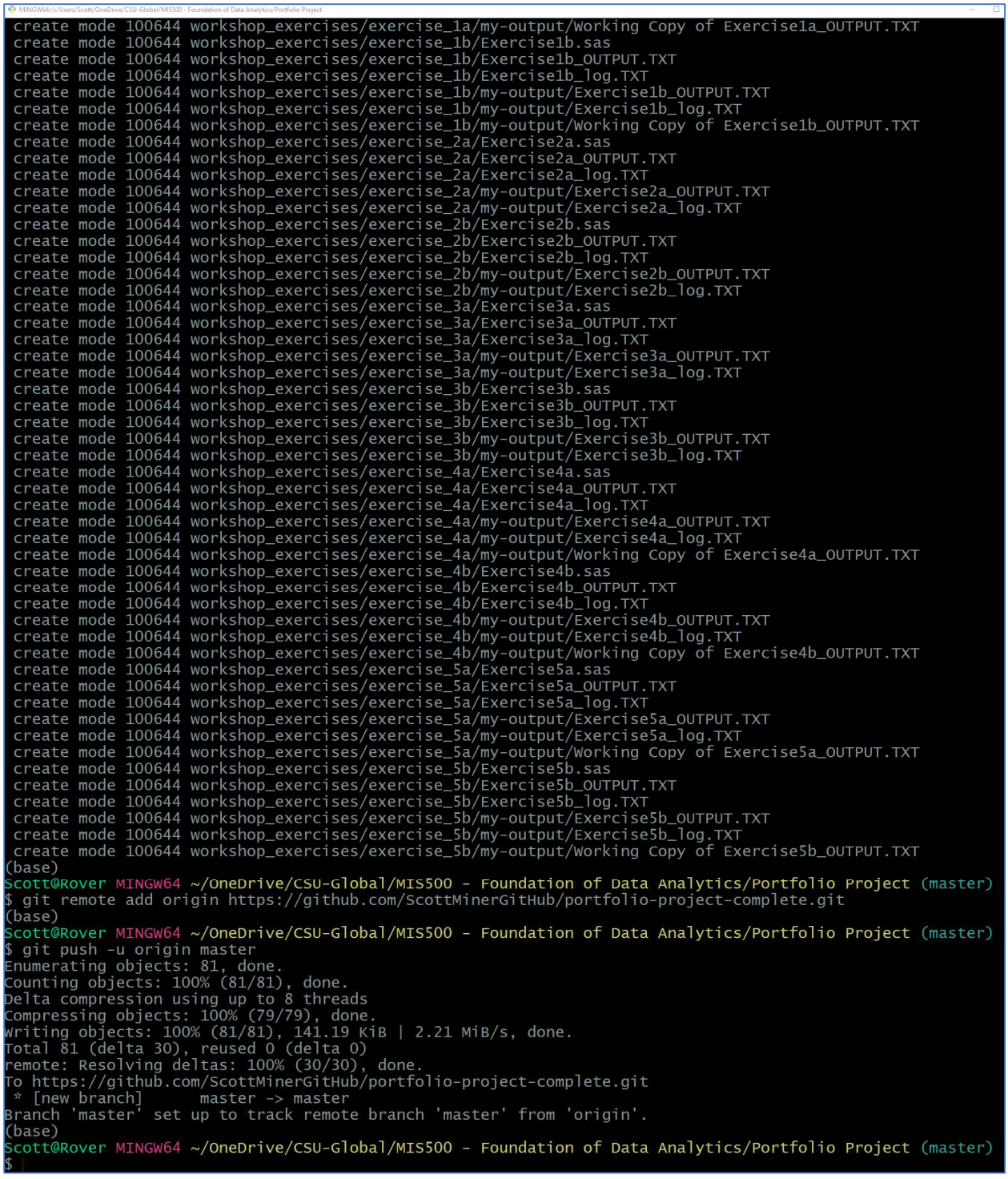

Figure 8. Initial screenshot of project being pushed to GitHub master

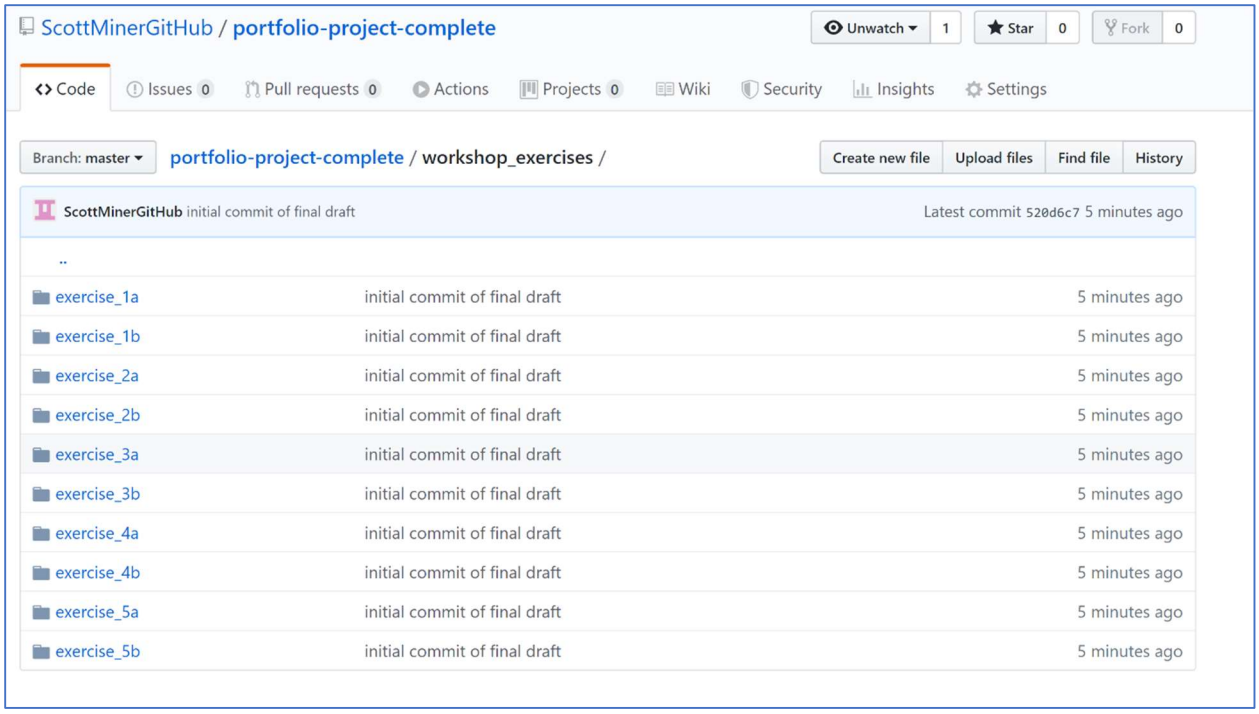

Figure 9. Project Repository on GitHub

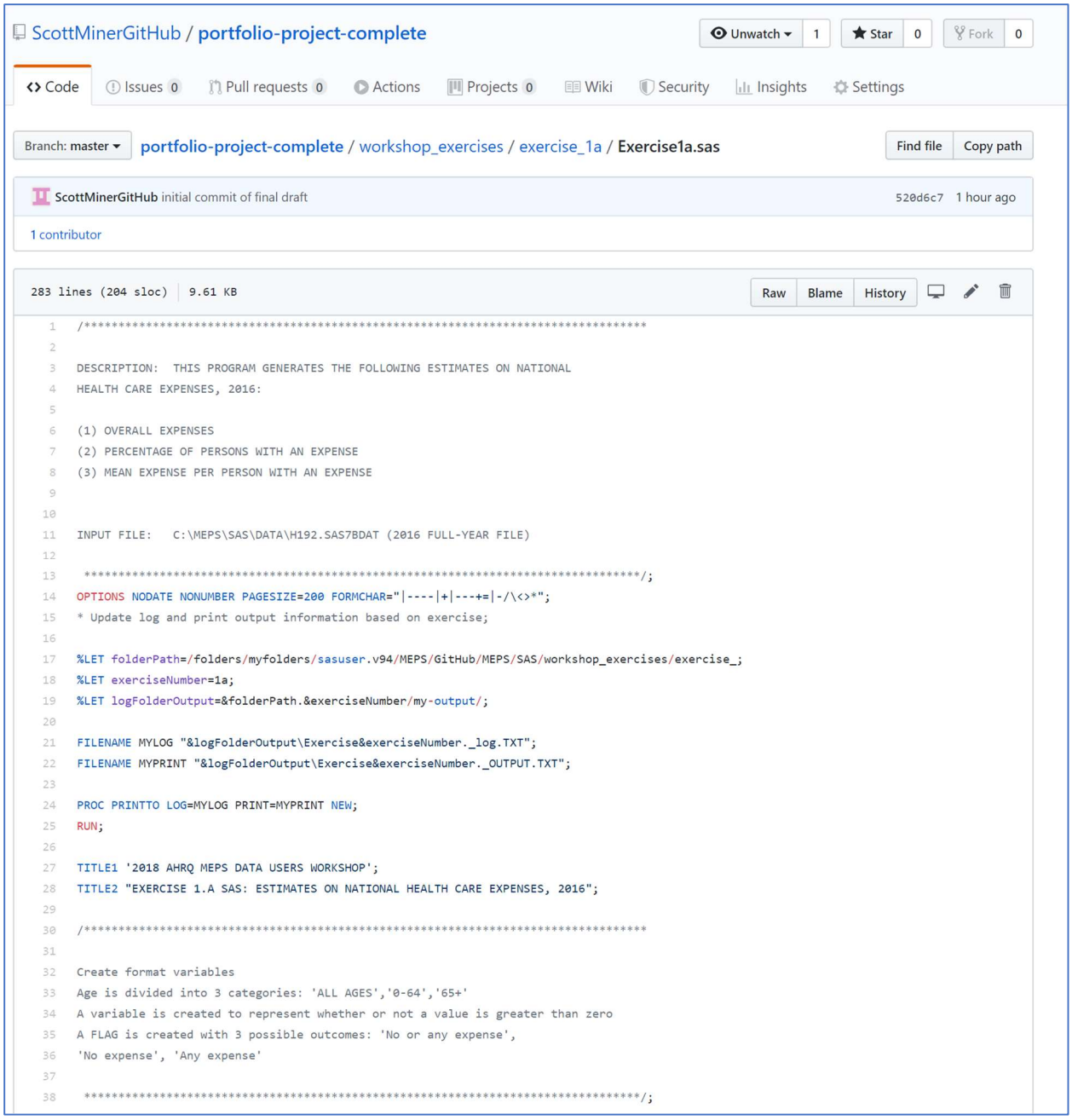

Figure 10. Exercise1a.SAS on GitHub with updated script comments

References

Elliott, A. C., & Woodward, W. A. (2016). *SAS essentials: Mastering SAS for data analytics* (Second edition). Hoboken, New Jersey: John Wiley and Sons, Inc.

HHS-AHRQ/MEPS. (n.d.). Retrieved December 1, 2019, from GitHub website: https://github.com/HHS-AHRQ/MEPS

Medical Expenditure Panel Survey PUF Codebook. (n.d.). Retrieved December 1, 2019, from https://meps.ahrq.gov/mepsweb/data_stats/download_data_files_codebook.jsp?PUFId=H181

MEPS HC-181 2015 Full Year Consolidated Data File. (n.d.). Retrieved November 30, 2019, from https://meps.ahrq.gov/data_stats/download_data/pufs/h181/h181doc.shtml

MEPS HC-192 2016 Full Year Consolidated Data File. (n.d.). Retrieved December 1, 2019, from https://meps.ahrq.gov/data_stats/download_data/pufs/h192/h192doc.shtml#Poverty2542

Moeller, J.F., Cohen S.B., & Hock E. (2002). Projecting National Medical Expenditure Survey data: a framework for MEPS projections. *MEPS Methodology Report No. 13.* AHRQ Pub. No. 02-0009.

PROC SURVEYMEANS: Missing Values: SAS/STAT(R) 9.2 User's Guide, Second Edition. (n.d.). Retrieved November 30, 2019, from https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_surveymeans_sect018.htm

ScottMinerGitHub—Overview. (n.d.). Retrieved November 30, 2019, from GitHub website: https://github.com/ScottMinerGitHub