

SAS CODE TO PRODUCE AGGREGATED VISIT STATISTICS AT THE PHYSICIAN OR FACILITY LEVEL

With the release of the 2005 National Ambulatory Medical Care Survey (NAMCS) and National Hospital Ambulatory Medical Care Survey (NHAMCS) public use micro-data files, provider weights became available for the first time. The variable PHYSWT on the NAMCS file allows researchers to make physician-level estimates. EDWT and OPDWT on the NHAMCS files allow researchers to make estimates at the emergency department level and the outpatient department level.

There are a number of variables on the files that pertain directly to characteristics of the physician or department, and basic instructions for using the provider weights with these variables can be found in the 2005 public use file documentation. However, there may be cases where researchers wish to link visit characteristics with providers and produce aggregated statistics at the provider level. This is a more complicated process, but can be accomplished. The code below shows one way to do this.

There are three basic steps to this process:

1. Organize the data in a DATA step, converting missing values for continuous variables to '.' and creating 0, 1 variables out of categorical variables where necessary. (If missing values for continuous variables are not converted, codes such as '9999' will be wrongly included in calculation of averages. For categorical variables used as analysis variables that have codes other than just 0 and 1, each must be converted to a separate 0/1 variable. See Example 1 below.)
2. Use PROC SUMMARY (or PROC MEANS) to create one record per provider along with the aggregate statistics for that provider.
3. Clean up the output file by converting proportions to percentages.

You are now ready to do further analysis as desired.

For questions, contact the Ambulatory Statistics Branch, 301-458-4600.

EXAMPLE 1

In this example, data from the 2005 ED public use file are being used to determine caseload percentages across emergency departments for expected sources of payment. (See Figure 7 in Advance Data Report No. 376 for a graphic display of similar results using 2003-04 data.) This example assumes you have downloaded the 2005 ED public use file and uncompressed it, downloaded the SAS input statement from the Ambulatory Health Care Data website, and have placed them both on your desktop in the folder 'c:\myfiles\nhamcs'.

```
FILENAME ed05 'c:\myfiles\nhamcs\ed2005';
FILENAME ed05inp 'c:\myfiles\nhamcs\ed05inp.txt';
FILENAME ed05for 'c:\myfiles\nhamcs\ed05for.txt';
FILENAME ed05lab 'c:\myfiles\nhamcs\ed05lab.txt';

%INC ed05for;

DATA edtest1;
INFILE ed05 MISSOVER LRECL=999;
%INC ed05inp;
%INC ed05lab;

/*The following IF statements use the hierarchical variable,
primary expected source of payment, to create separate variables
for each category of interest.*/

IF paytype=1 THEN privpay=1; ELSE privpay=0;
IF paytype=2 THEN mcarepay=1; ELSE mcarepay=0;
IF paytype=3 THEN mcaidpay=1; ELSE mcaidpay=0;

/*Note that in the hierarchical schema on which paytype is
based, Medicaid takes highest precedence among multiple
expected payment sources. If you wish to develop a different
priority, you can use the multiple expected pay source categories
PAYPRIV, PAYMCARE, PAYMCAID, PAYSELF, etc. on the 2005 public use
file to do so.*/

PROC SUMMARY DATA=edtest1 NWAY;
/*You could also use PROC MEANS; these are "sister" procedures
in SAS and are nearly identical. Note that if you choose to use PROC
MEANS, the default option does not create an output dataset as
PROC SUMMARY does. You can, however, create one using an OUTPUT
Statement. See a SAS Language or Procedures manual for more
information. The NWAY option suppresses marginal counts.*/

CLASS hospcode;
/*CLASS groups records by the hospital identifier HOSPCODE. Each
hospital on the file only has one emergency department, so the
results reflect unique emergency departments.*/
```

```

ID edwt cstratm cpsum;
/*This is where you can retain any variables you wish to keep in your
dataset that do NOT vary by HOSPCODE. Be sure to select only variables
that have the same value for each visit record within a particular
HOSPCODE! (NOTE: EDWT only occurs on the first record for each HOSPCODE
and is missing for other records, but because PROC SUMMARY concatenates
all of the ID variables and takes the maximum value of the concatenated
string as the source for the new record values, this will give
appropriate results.)*/

VAR privpay mcarepay mcaidpay uninspay;
/*These are the variables that you wish to aggregate within each
hospital.*/

OUTPUT OUT=edtest2 MEAN=pctpriv pctmcare pctmcaid pctunins;
/*This assigns a new output file name and tells SAS what kind of
aggregated statistics you want for each variable from the VAR
statement. It also renames the new aggregated variables into something
more meaningful.*/

PROC PRINT DATA=edtest2 (OBS=10); RUN;
/*The PROC PRINT statement gives you a sample of your output
Dataset so you can make sure it looks the way you want it to.*/

DATA edpay; SET edtest2;
pctpriv=pctpriv*100;
pctmcare=pctmcare*100;
pctmcaid=pctmcaid*100;
pctunins=pctunins*100;
/*Converts proportions to percentages.*/

PROC UNIVARIATE DATA=edpay;
VAR pctpriv pctmcare pctmcaid pctunins;

/*This shows you the distribution across hospitals of percentages of
visits by each paytype.*/

```

EXAMPLE 2

In this example, 2005 data are used to produce average waiting time to see a physician in emergency departments by MSA location (i.e. whether the emergency department is in a metropolitan statistical area or not).

```

FILENAME ed05 'c:\myfiles\nhamcs\ed2005';
FILENAME ed05inp 'c:\myfiles\nhamcs\ed05inp.txt';
FILENAME ed05for 'c:\myfiles\nhamcs\ed05for.txt';
FILENAME ed05lab 'c:\myfiles\nhamcs\ed05lab.txt';

%INC ed05for;

DATA edwait1;
INFILE ed05 MISSOVER LRECL=999;
%INC ed05inp;
%INC ed05lab;

```

```
IF waittime = 9999 THEN waittime = . ;
/*Recodes missing values on continuous variables to . */

RUN;

PROC SUMMARY DATA=edwait1 NWAY;
CLASS hospcode; ID edwt msa cstratm cpsum;
VAR waittime;

OUTPUT OUT=edwait2 MEAN=avgwaittime;
/*Sends output to new temporary SAS dataset and
renames waittime to reflect conversion to average.*/

PROC PRINT DATA=edwait2 (obs=10); RUN;

PROC UNIVARIATE DATA=edwait2;
CLASS msa;
VAR avgwaittime; WEIGHT edwt;
TITLE 'Distribution of average waiting times in hospital emergency
departments in MSAs and non-MSAs';
.

/*Document created 8/21/12*/
```