

SAS User File for H220IF2 Data

This file contains information and sample SAS programs to create a permanent SAS dataset for users who want to use SAS in processing the MEPS data provided in this PUF release.

There are two ways to create a permanent SAS dataset, using either the SAS transport data file (H220IF2.SSP) or the ASCII data file (H220IF2.DAT) supplied in this PUF release. Section A provides a sample SAS program for the first alternative, which is to convert the SAS transport data file to a regular SAS dataset using the SAS procedure: CIMPORT. Section B provides a sample SAS program for the second alternative, which is to read data from the ASCII data file using a SAS DATA step with INFILE, INPUT, and LABEL statements. Section C explains format-related SAS statements that a user may optionally use when working with the SAS dataset. Examples of SAS programs (DATA step or PROC) are provided in all three sections, primarily for the benefit of inexperienced users. Section D contains complete SAS statements that must be used in the programs described in Sections B and C.

\*\*\*\*\*

The sample SAS programs provided in Sections A and B show how to create a permanent SAS dataset from the data files provided in this PUF release.

A. A Sample SAS Program for Converting the SAS Transport File to a Permanent SAS Dataset

The SAS procedure CIMPORT will read a SAS transport file and convert the data to regular SAS format, storing the output in a permanent SAS dataset. This permanent SAS dataset can then be used for all future processing and analyses.

Below is a sample SAS program that can be used to convert the SAS transport file to a permanent SAS dataset (in a Windows environment).

```
LIBNAME PUFLIB 'C:\MEPS\SASDATA';
FILENAME IN1 'C:\MEPS\DOWNLOAD\H220IF2.SSP';

PROC CIMPORT DATA=PUFLIB.H220IF2 INFILE=IN1;
RUN;
```

Below are SAS statements to print a list of variables and a few sample records from the permanent SAS dataset:

```
PROC CONTENTS DATA=PUFLIB.H220IF2;
TITLE 'List of Variables in MEPS H220IF2 SAS Dataset';
RUN;

PROC PRINT DATA=PUFLIB.H220IF2 (OBS=20);
TITLE 'First 20 Observations in MEPS H220IF2 SAS Dataset';
RUN;
```

The LIBNAME statement tells SAS the location (directory name) to store the permanent SAS dataset which is output by PROC CIMPORT. The FILENAME statement tells SAS the location (complete directory and file name) of the input SAS transport data file.

NOTES:

- 1) The names used in the LIBNAME and FILENAME statements shown above (i.e., PUFLIB, IN1) are arbitrary; they are only temporary aliases.
- 2) The directory and file names used in the LIBNAME and FILENAME statements shown above are Windows syntax and may need to be modified for other operating systems such as UNIX and Linux.
- 3) H220IF2 is the internal SAS dataset name (also the PC file name, without the extension) prior to the creation of the SAS transport data file. After running PROC CIMPORT, the output SAS dataset assumes the same dataset name (or file name). Hence, in the example above, a file named H220IF2.SAS7BDAT will be created under the C:\MEPS\SASDATA directory when PROC CIMPORT runs successfully.
- 4) The SAS transport file H220IF2.SSP was created from a SAS V9 data file, using PROC CPORT. This file should work with earlier versions of SAS, although it has not been tested with those versions. Users who are unable to use this SAS transport file should instead convert the ASCII data file H220IF2.DAT to a SAS dataset as described in Section B.

B. A Sample SAS Program for Converting the ASCII Data File to a Permanent SAS Dataset

The complete SAS statements (INPUT and LABEL) included in Section D are intended to save time for those users wishing to create a permanent SAS dataset from the H220IF2.DAT ASCII data file. These statements must be used in combination with other SAS statements to create the appropriate SAS program, as shown below. To use the statements provided in Section D to create a SAS program, you will need an ASCII text editor. If you are using an interactive form of SAS (Windows, UNIX, Linux, etc.), use the editor provided as part of the SAS software.

Following is a sample Windows SAS program that will convert the ASCII data file to SAS format:

```
LIBNAME PUFLIB 'C:\MEPS\SASDATA';
FILENAME IN1 'C:\MEPS\DOWNLOAD\H220IF2.DAT';

DATA PUFLIB.H220IF2;
INFILE IN1 LRECL=77;
INPUT .....; * to user: insert the complete INPUT statement that is
provided in Section D;
LABEL .....; * to user: insert the complete LABEL statement that is
provided in Section D;
RUN;
```

Here is an explanation of the SAS statements used in the program above.

LIBNAME statement: This tells SAS the location (directory name) of the permanent SAS dataset.

FILENAME statement: This tells SAS the location of the input ASCII data file.

DATA statement: This signifies the beginning of a SAS DATA step and specifies the output SAS dataset, referencing the LIBNAME entry (PUFLIB) and assigning an internal SAS dataset name (H220IF2). In the example, after the successful completion of the DATA step, a PC file named H220IF2.SAS7BDAT would have been created in the C:\MEPS\SASDATA directory.

INFILE statement: This tells SAS the location (directory and file name) of the input ASCII data file. Also provided is the logical record length (77 bytes), with the default of RECFM=V implied when this parameter is omitted. LRECL and RECFM are optional parameters in the INFILE statement. With regard to these options, please note the following: the ASCII data file H220IF2.DAT contains a 2-byte carriage return/line feed at the end of each record. When converting to a PC-SAS file, the LRECL option should be used to specify the record length to avoid use of a default record length by PC-SAS. If the RECFM=V option is used, the LRECL option must be specified as the logical record length (e.g., 77 for H220IF2.DAT). If RECFM=F is used, then the LRECL value must be specified as the logical record length plus 2 (79 for H220IF2.DAT). Note that if the RECFM option is omitted, then the default option of RECFM=V is automatically used, and LRECL should be specified as the logical record (77 for H220IF2.DAT).

INPUT statement: This specifies the input record layout, giving names, the beginning column positions, and the lengths for data items (which become SAS variables) in the ASCII data file (H220IF2.DAT). Variable type (numeric or character) is also defined by character variables having a dollar sign (\$) directly before the variables length, while numeric variables do not have a dollar sign.

LABEL statement: This associates descriptive names with the SAS variables.

RUN statement: This tells SAS to execute all commands up to this point.

C. Optional Format-related SAS Statements

If a user wants to use formats for the SAS variables, a SAS format catalog must first be created. Below is a SAS program that will accomplish this:

```
LIBNAME PUFLIB 'C:\MEPS\SASDATA';

PROC FORMAT LIBRARY=PUFLIB;
VALUE .....; * to user: insert the complete set of VALUE statements found
in Section D;
VALUE .....;
.....;
RUN;
```

Below is an example of how to use the SAS formats defined by the PROC FORMAT procedure:

```
LIBNAME PUFLIB 'C:\MEPS\SASDATA';
OPTIONS FMTSEARCH=(PUFLIB);

PROC FREQ DATA=PUFLIB.H220IF2;
TABLES .... / LIST MISSING;
FORMAT varnam1 fmtnam1. Varnam2 fmtnam2. ....;
* to user: substitute varnam1 and fmtnam1 with actual variable names and
format names;
* Insert the FORMAT statement provided in Section D, if you are using
all the variables in the TABLES statement;
TITLE 'Frequency Distributions .....';
RUN;
```

Here is an explanation of the SAS statements used above.

LIBNAME statement: This tells SAS the location (directory name) of the SAS format library. Please note that SAS datasets (file name extension is 'SAS7BDAT') and format

catalog (file name extension is 'SAS7BCAT') can be stored under the same directory.

OPTIONS FMTSEARCH=...: This specifies the SAS format library.

PROC FORMAT statement: This identifies the SAS procedure that will make SAS formats according to VALUE statements. Formats will be stored in a file named FORMATS.SAS7BCAT. Please note that the option 'LIBRARY=...' can be omitted if the user does not want to create a permanent SAS format library. When simply 'PROC FORMAT;' is used, the formats are defined only for the duration of the batch SAS program or an interactive SAS session.

VALUE statement: This gives a) names to formats; and b) descriptive labels for individual values, or range of values. The format names can then be invoked using a FORMAT statement if desired.

PROC FREQ statement: This identifies the SAS procedure that generates frequency distributions of variables specified in the TABLES statement, formatted if a FORMAT statement is used. The input SAS dataset is specified in the 'DATA=' option.

FORMAT statement: This associates existing formats with variables. When using this statement, the formats must have already been created with a PROC FORMAT procedure.

RUN statement: This tells SAS to execute all commands up to this point.

NOTES:

- 1) Use of formats is entirely optional, and depends on the types of analyses that you are doing. It is recommended that you create and use them as appropriate.
- 2) The names used in the LIBNAME and FILENAME statements shown above (i.e., PUFLIB, IN1) are arbitrary; they are only temporary aliases.
- 3) The file and directory specifications in the LIBNAME and FILENAME statements are Windows syntax and may need to be modified for other operating systems such as UNIX and Linux.

D. SAS Statements

This section contains SAS INPUT, LABEL, FORMAT, and VALUE statements for use in converting the ASCII H220IF2.DAT file into a SAS dataset, and for creating SAS formats.

```
* INPUT STATEMENTS;
INFILE IN LRECL=77;

INPUT @1          DUPERSID          $10.0
      @11         RXLKIDX           $32.0
      @43         EVNTIDX           $16.0
      @59         LINKIDX           $16.0
      @75         EVENTYPE          1.0
      @76         PANEL             2.0
;

* FORMAT STATEMENTS;
FORMAT DUPERSID          $DUPERS.
      RXLKIDX            $RXLKIDX.
      EVNTIDX            $EVNTIDX.
      LINKIDX            $LINKIDX.
      EVENTYPE           $EVENTYPE.
      PANEL              $PANEL.
;

* LABEL STATEMENTS;
LABEL DUPERSID          ="PERSON ID (DUID + PID)"
      RXLKIDX            ="RECORD ID: EVNTIDX + LINKIDX"
      EVNTIDX            ="EVENT ID"
      LINKIDX            ="ID FOR LINKAGE TO COND/OTH EVENT FILES"
      EVENTYPE           ="TYPE OF EVENT RX IS LINKED TO"
      PANEL              ="PANEL NUMBER"
;

* VALUE STATEMENTS;
VALUE $DUPERS
  '2320018102' - '2579869101' = "VALID ID"
;
VALUE $EVENTYPE
  1 = "1 OFFICE-BASED MEDICAL PROVIDER VISIT"
  2 = "2 OUTPATIENT DEPARTMENT VISIT"
  3 = "3 EMERGENCY ROOM VISIT"
  4 = "4 INPATIENT HOSPITAL STAY"
;
VALUE $EVNTIDX
  '2320018102209401' - '2579869101000601' = "VALID ID"
;
VALUE $LINKIDX
  '2320018102001703' - '2579869101002303' = "VALID ID"
;
VALUE $PANEL
  23 = "23 PANEL 23"
  24 = "24 PANEL 24"
  25 = "25 PANEL 25"
;
VALUE $RXLKIDX
  '23200181022094012320018102007603' - '25798691010006012579869101002303' = "VALID ID"
;
;
```