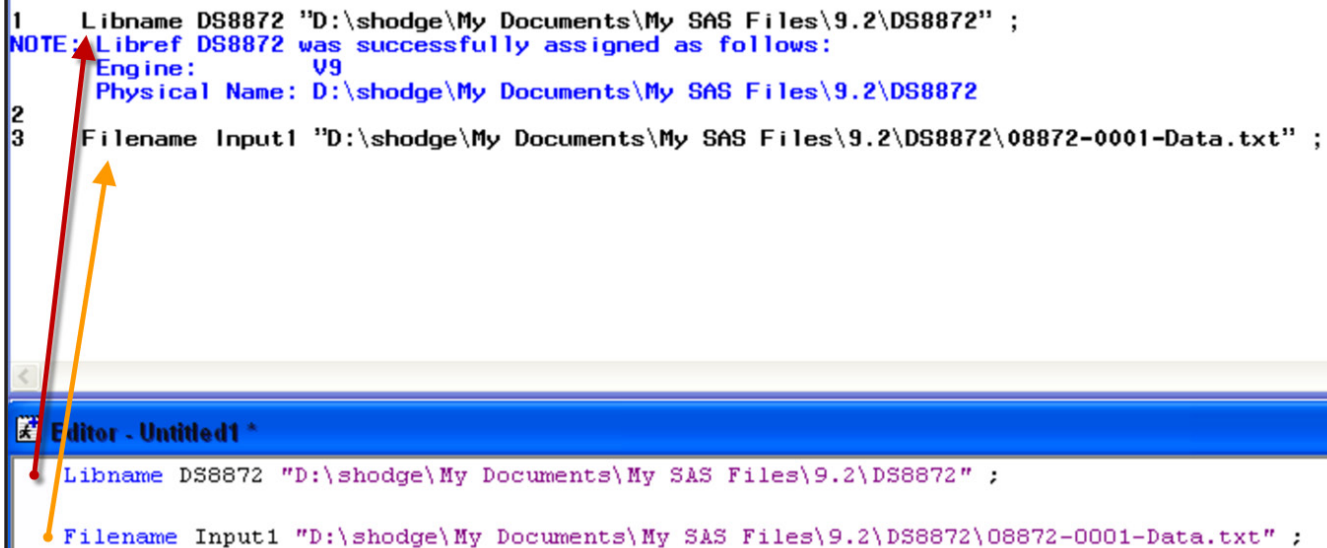


Using an ICPSR set-up file to create a SAS dataset

Name library and raw data files.



```
1 Libname DS8872 "D:\shodge\My Documents\My SAS Files\9.2\DS8872" ;  
NOTE: Libref DS8872 was successfully assigned as follows:  
Engine: V9  
Physical Name: D:\shodge\My Documents\My SAS Files\9.2\DS8872  
2  
3 Filename Input1 "D:\shodge\My Documents\My SAS Files\9.2\DS8872\08872-0001-Data.txt" ;
```

Editor - Untitled1 *

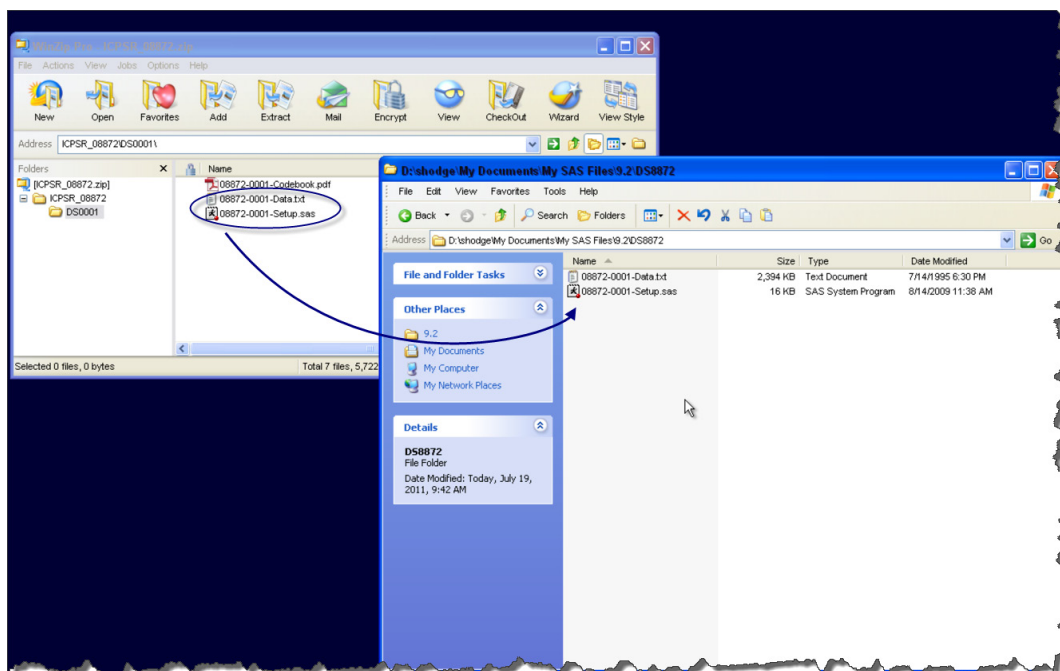
```
Libname DS8872 "D:\shodge\My Documents\My SAS Files\9.2\DS8872" ;  
Filename Input1 "D:\shodge\My Documents\My SAS Files\9.2\DS8872\08872-0001-Data.txt" ;
```

From the Start menu, launch SAS, and in the Editor program, write the codes to create and name a folder in the SAS permanent file for this project, "DS8872," and rename the raw data file, "Input1," and save in the new folder, DS8872. Be sure to observe the SAS filenames convention, which states that names are

- no longer than 8 characters,
- start with a letter, and
- contain only letters, numbers, or underscores (_).

The Log above the Editor confirms action complete and shows errors, if any.

Move data and set-up file into new SAS folder.



From the zip folder downloaded from ICPSR, select the "set-up.sas" file and the "data.txt" file and drag and drop them into the new folder, DS8872. If there are more than one dataset in the downloaded file, be sure that the set-up file and the data file have the same dataset number, which is the second set of numbers in the name, such as ".0001."

Delete comment marks from procedures.

```

08872.0001-Setup.sas
| setup, users should remove the SAS comment indicators from the desired
| section(s).
|
|-----*/
* SAS PROC FORMAT;
/*
PROC FORMAT;
  VALUE cgidmont 00='(00) UNKNOWN' 01='(01) JANUARY' 02='(02) FEBRUARY' 03='(03) MARCH'
  04='(04) APRIL' 05='(05) MAY' 06='(06) JUNE' 07='(07) JULY' 08='(08) AUGUST'
  09='(09) SEPTEMBER' 10='(10) OCTOBER' 11='(11) NOVEMBER' 12='(12) DECEMBER' ;
  VALUE cgiddayw 00='(00) UNKNOWN' 99='(99) UNKNOWN' ;
  VALUE monthfff 00='(00) UNKNOWN' 01='(01) JANUARY' 02='(02) FEBRUARY' 03='(03) MARCH'
  04='(04) APRIL' 05='(05) MAY' 06='(06) JUNE' 07='(07) JULY' 08='(08) AUGUST'
  09='(09) SEPTEMBER' 10='(10) OCTOBER' 11='(11) NOVEMBER' 12='(12) DECEMBER' ;
  VALUE wday 0='(0) UNKNOWN' 1='(1) SUNDAY' 2='(2) MONDAY' 3='(3) TUESDAY' 4='(4) WEDNESDAY'
  5='(5) THURSDAY' 6='(6) FRIDAY' 7='(7) SATURDAY' 99='(99) UNKNOWN' ;
  VALUE aday 0='(0) EXACT' 1='(1) WITHIN 1 WEEK' 2='(2) WITHIN 2 WEEKS'
  3='(3) WITHIN 1 MONTH' 4='(4) WITHIN 2 MONTHS' 5='(5) WITHIN 3 MONTHS' ;
  VALUE durg -1='(-1) UNKNOWN' 0='(0) ONE DAY OR LESS' 1='(1) UP TO 1 WEEK'
  2='(2) UP TO 2 WEEKS' 3='(3) UP TO 1 MONTH' 4='(4) UP TO 2 MONTHS'
  5='(5) UP TO 3 MONTHS' ;
  VALUE durd -1='(-1) UNKNOWN' ;
  VALUE durgues 0='(0) EXACT' 1='(1) GUESS' ;
  VALUE durh -1='(-1) UNKNOWN' ;
  VALUE durgues 0='(0) EXACT' 1='(1) GUESS' ;
  VALUE gtyp 1='(1) VIOLENCE' 2='(2) MEETING' 3='(3) GATHERING' 4='(4) DELEGATION'
  99='(99) OTHER' ;
  VALUE type 0='(0) UNKNOWN' 1='(1) POACHERS VS. GAMEKEEPERS'
  2='(2) SMUGGLERS VS. CUSTOMS' 3='(3) BRAWLS IN DRINKING PLACES'
  4='(4) OTHER VIOLENT GATHERINGS' 5='(5) ATTACKS ON BLACKLEGS'

```

Comment marks precede (/*) and follow (*/) statements that are not enabled in the set-up file. These statements contain the labels and names for values and variables, labeling for missing data and other useful information. To include the statements, delete the marks around the statement. Continue working through the file until you reach the DATA and INFILE statements.

Edit Data and Infile Commands.

```

VALUE pday -1='(-1) UNKNOWN' 0='(0) EVNT LASTED < 24 HOURS' ;
VALUE pdaymoe -1='(-1) UNKNOWN' ;
VALUE phrs -1='(-1) UNKNOWN' ;
VALUE phrsmoe -1='(-1) UNKNOWN' ;
VALUE arde -1='(-1) UNKNOWN' ;
VALUE ardemoe -1='(-1) UNKNOWN' ;
VALUE arae -1='(-1) UNKNOWN' ;
VALUE araemoe -1='(-1) UNKNOWN' ;
VALUE woun -1='(-1) UNKNOWN' ;
VALUE wounmoe -1='(-1) UNKNOWN' ;
VALUE kill -1='(-1) UNKNOWN' ;
VALUE killmoe -1='(-1) UNKNOWN' ;
VALUE CL EVENT 2='(2) GBS NON-SWING EVENT' 3='(3) GBS SWING EVENT' ;

* SAS DATA, INFILE, INPUT STATEMENTS;

DATA DS8872.da8872;
INFILE Input1 LRECL=302;
INPUT

```

The DATA statement names the new dataset preceded by the library's name, "DS8872.da8872." The INFILE statement points to the raw data or its location to create the new dataset. We have already given the raw data a SAS filename, "Infile1." Highlight and delete "Data-filename" and replace it with "Infile1." Retain the LRECL if given.

Continue working through the file removing comment marks, as appropriate.

Run set-up file.

```

08872-0001.Setup.sas *
IF {ARAE = -1} THEN ARAE = .;
IF {ARAEMOE = -1} THEN ARAEMOE = .;
IF {WOUN = -1} THEN WOUN = .;
IF {WOUNMOE = -1} THEN WOUNMOE = .;
IF {KILL = -1} THEN KILL = .;
IF {KILLMOE = -1} THEN KILLMOE = .;

* SAS FORMAT STATEMENT;

FORMAT CGIDMONT cgidmont. CGIDDAYW cgiddayw. MONTH monthfff.
        WDAY wday. ADAY aday. DURG durg.
        DURD durd. DURDGUES durdgues. DURH durh.
        DURHGUES durhgues. GTYP gtyp. TYPE type.
        COUNTY county. MC mc. LONDTMS londtms.
        GM gm. AR ar. HPD hpd.
        MOP mop. VAP vap. LC lc.
        PLOW plow. PHI phi. PNUM pnum.
        PNUMCALC pnumcalc. PDAY pday. PDAYMOE pdaymoe.
        PHRS phrs. PHRSMOE phrsmoe. ARDE arde.
        ARDEMOE ardemoe. ARAE arae. ARAEMOE araemoe.
        WOUN woun. WOUNMOE wounmoe. KILL kill.
        KILLMOE killmoe. SET set. ;

RUN ;

```

Once the DATA and INFILE commands have been edited to include the names and location of the files and the naming of variables and value labels have been enabled, check to be sure that there is a RUN; command in place. Then click on the icon at the top of the window to run the file.

Confirm success.

```

314      PHRS phrs. PHRSMOE phrsmoe. ARDE arde.
315      ARDEMOE ardemoe. ARAE arae. ARAEMOE araemoe.
316      WOUN woun. WOUNMOE wounmoe. KILL kill.
317      KILLMOE killmoe. SET set. ;
318
319
320 RUN ;

NOTE: The infile INPUT1 is:
      Filename=D:\shodge\My Documents\My SAS Files\9.2\DS8872\08872-0001-Data.txt,
      RECFM=V,LRECL=302,File Size (bytes)=2450664,
      Last Modified=14Jul1995:18:30:12,
      Create Time=14Jul1995:18:30:12

NOTE: 8088 records were read from the infile INPUT1.
      The minimum record length was 302.
      The maximum record length was 302.

NOTE: The data set DS8872.DA8872 has 8088 observations and 66 variables.

NOTE: DATA statement used (Total process time):
      real time          19.90 seconds
      cpu time           0.18 seconds

```

Here is the confirmation in the Log file that the new SAS dataset has been created.

Edit and save set-up file.

```

FORMAT: associates the formats created by the PROC FORMAT step with
the variables named in the INPUT statement.

NOTE: Users should modify this setup file to suit their specific needs.
Sections for PROC FORMAT, FORMAT, and MISSING VALUE RECODES have been
commented out (i.e., '/*'). To include these sections in the final SAS
setup, users should remove the SAS comment indicators from the desired
section(s).

-----*/

*Create library and name raw data file ;

Libname DS8872 "D:\shodge\My Documents\My SAS Files\9.2\DS8872" ;

Filename Input1 "D:\shodge\My Documents\My SAS Files\9.2\DS8872\08872-0001-Data.txt" ;

* SAS PROC FORMAT;

PROC FORMAT;
  VALUE cgldmont 00=(00) UNKNOWN' 01=(01) JANUARY' 02=(02) FEBRUARY' 03=(03) MARCH'
  04=(04) APRIL' 05=(05) MAY' 06=(06) JUNE' 07=(07) JULY' 08=(08) AUGUST'
  09=(09) SEPTEMBER' 10=(10) OCTOBER' 11=(11) NOVEMBER' 12=(12) DECEMBER' ;
  VALUE cglddayw 00=(00) UNKNOWN' 99=(99) UNKNOWN' ;
  VALUE monthfff 00=(00) UNKNOWN' 01=(01) JANUARY' 02=(02) FEBRUARY' 03=(03) MARCH'
  04=(04) APRIL' 05=(05) MAY' 06=(06) JUNE' 07=(07) JULY' 08=(08) AUGUST'
  09=(09) SEPTEMBER' 10=(10) OCTOBER' 11=(11) NOVEMBER' 12=(12) DECEMBER' ;
  VALUE wday 0=(0) UNKNOWN' 1=(1) SUNDAY' 2=(2) MONDAY' 3=(3) TUESDAY' 4=(4) WEDNESDAY'
  5=(5) THURSDAY' 6=(6) FRIDAY' 7=(7) SATURDAY' 99=(99) UNKNOWN' ;
  VALUE aday 0=(0) EXACT' 1=(1) WITHIN 1 WEEK' 2=(2) WITHIN 2 WEEKS' ;

```

Copy and paste the Libname and Filename commands from the beginning of this process into the beginning of the set-up file after the heading and before the first PROC FORMAT statement. Write a short statement about what the statements accomplish preceded by an asterisk (*) and followed by a semi-colon (;). Save the file to the folder, DS8872 with a name that is descriptive, such as “ed-ited.08872-0001.setup.”

Create Formats Catalog for future use.

```

WOUN woun. WOUNMOE wounmoe. KILL kill.
KILLMOE killmoe. SET set. ;

Run;

*Write formats to DS8872;

Proc Format cntlout = DS8872.Formats ;

Run ;

```

Start with a short statement about the statement preceded by an asterisk (*) and followed by a semi-colon (;). This statement saves the Formats Catalog that the set-up file created so it can be used with the dataset in the future. Use a PROC FORMAT statement and the option “CNTLOUT=DS8872.Formats” to create the catalog and add a “Run ;” command. Be sure to close each line with a semi-colon. Highlight the PROC FORMAT AND RUN statements and go to the menus at the top and click on “Submit.” After confirming in the log file that the catalog has been successfully created, save the edited set up file again. In subsequent sessions using this dataset, use a PROC FORMAT and CNTLIN option to make the data and variable labels and names available in the SAS session.

Open dataset for subsequent use.

```

Editor - Untitled1 *

*load dataset into SAS including formats ;

Libname DS8872 "D:\shodge\My Documents\My SAS Files\9.2\DS8872" ;
title "DS8872.DA8872" ;
title "DS8872.Formats";
proc format cntlin=DS8872.Formats ;
proc contents data=DS8872.DA8872 ;

run ;

```

Open SAS and begin in a new Program Editor. Write a short note about opening a dataset in SAS and applying the formats catalog. Remember to start with an asterisk (*) and end with a semi-colon (;). Name the library DS8872 using the "LIBNAME" statement and enter the path to the file. Titling the files is optional but is included here following the LIBNAME. Use the PROC FORMAT statement with the option CNTLIN and point to the formats catalog to make the formats accessible to SAS, "cntlin=ds8872.formats." A PROC CONTENT statement that describes the dataset can be used but is also optional. Add a "RUN ;" command at the end of the script and go to the menus at the top and click on "Submit."

Confirm dataset opened in SAS.

Output - (Untitled)

DS8872.Formats09:55 Wednesday, July 20, 2011

The CONTENTS Procedure

Data Set Name	DS8872.DA8872	Observations	8088
Member Type	DATA	Variables	66
Engine	V9	Indexes	0
Created	Wednesday, July 20, 2011 08:41:30 AM	Observation Length	552
Last Modified	Wednesday, July 20, 2011 08:41:30 AM	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	WINDOWS_32		
Encoding	wlatin1 Western (Windows)		

Engine/Host Dependent Information

Data Set Page Size	16384
Number of Data Set Pages	280
First Data Page	1
Max Obs per Page	29
Obs in First Data Page	12
Number of Data Set Repairs	0
Filename	D:\shodge\My Documents\My SAS Files\9.2\DS8872\da8872.sas7bdat
Release Created	9.0202M0
Host Created	XP_PRO

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Label
9	ADAY	Num	8	ADAY.	ACCURACY OF STARTING DATE
26	AR	Num	8	AR.	SOURCE=ANNUAL REGISTER
42	ARAE	Num	8	ARAE.	NUMBER OF PERSONS ARRESTED AFTER EVENT
43	ARAE MOE	Num	8	ARAE MOE.	MARGIN OF ERROR-ARRESTED AFTER EVENT
40	ARDE	Num	8	ARDE.	NUMBER OF PERSONS ARRESTED DURING EVENT
44	ARDE MOE	Num	8	ARDE MOE.	MARGIN OF ERROR-ARRESTED DURING EVENT

The results of the PROC CONTENT statement confirms the location of the file, the number of cases, and the path to the file. This dataset is ready for analysis.