22S:166
Computing in Statistics

Other Software Packages
Proc import
A bit on SAS macro language

Lecture 25
Nov. 19, 2008

Kate Cowles
374 SH, 335-0727
kcowles@stat.uiowa.edu

Other software packages

- Microsoft Excel
  - spreadsheet
  - very convenient for entering data in flat-file format
  - clients very frequently bring data to statisticians in Excel format
  - NOT reliable and accurate for doing statistical analysis
- Microsoft Access
  - relational database management system

Reading data files into SAS from other software packages

- Import Wizard
  - point-and-click interactive reading
  - convenient if file only needs to be read once
  - can write `proc import` code to be copied into programs
- `proc import`
  - can be used instead of data step in SAS programs
  - much more convenient if file needs to be read in multiple programs, or program using file needs to be run repeatedly

Importing from Other Sources

Types of files that the Import Wizard and/or `proc import` can read

Identifier	Input Data Source	Extension
ACCESS	Microsoft Access database	.MDB
DBF	dBASE file	.DBF
WK1	Lotus 1 spreadsheet	.WK1
WK3	Lotus 3 spreadsheet	.WK3
WK4	Lotus 4 spreadsheet	.WK4
EXCEL	Excel V 4 or 5 spreadsheet	.XLS
EXCEL4	Excel V 4 spreadsheet	.XLS
EXCEL5	Excel V 5 spreadsheet	.XLS
EXCEL97	Excel 97 spreadsheet	.XLS
DLM	delimited file (default is blank)	.*
CSV	delimited file (comma-sep vals)	.CSV
TAB	delimited file (tab-delimited)	.TXT

Restriction: The data sources available to you depend on the SAS/ACCESS product s that you have licensed. If you do not have any SAS/ACCESS products licensed, then the only types of data source files available to you are .CSV, .TXT, and d elimited files.

Example

- from R or Splus
  - use `write.table` to write data out as a delimited file

```
Data frame that comes with R

> USArrests
             Murder Assault UrbanPop Rape
Alabama        13.2     236       58 21.2
Alaska         10.0     263       48 44.5
Arizona         8.1     294       80 31.0
Arkansas        8.8     190       50 19.5
California      9.0     276       91 40.6
.
.
.
```

R command to write out file as tab-delimited data file

```
> write.table( USArrests, file="C:\\My Documents\\166\\USArrests.txt",
  sep="\t", quote = FALSE, col.names=TRUE)
```

Now in SAS....

```
File / Import Data

Import Wizard
Select a data source from the list below

Choose "Delimited File (*.*)"

Where is the file located?
Give full path name, e.g.
C:\My Documents\166\USArrests.txt


 Choose SAS destination:
Library:  (defaults to WORK)
Member:  (fill in name of your choice; e.g. USArrest)


 Question as to whether you want wizard to generate proc
import statements so you can just run them next time
```

What it generated

```
PROC IMPORT OUT= WORK.usarrest
            DATAFILE= "C:\My Documents\166\USArrests.txt"
            DBMS=DLM REPLACE;
     DELIMITER='00'x;  * needed correction to DELMITED='09'x ;
     GETNAMES=YES;
     DATAROW=2;
RUN;
```

Example of reading Access database

```
PROC IMPORT OUT= WORK.courses
            DATATABLE= "Courses"
            DBMS=ACCESS97 REPLACE;
     DATABASE="c:\my documents\166\univ0_v7";
RUN;
```

Overview of SAS Macro Programming

- purpose is to make SAS programming more efficient and to reduce coding errors
- macro variables
  - enable substitution of text into SAS programs
- macro programs
  - enable performing the same task on different inputs without rewriting code

Example dataset

Data Set Name: BOOKS.YTDSALES — Observations: 695
Member Type: DATA — Variables: 10
Engine: V8 — Indexes: 0
Created: 7:36 Friday, October 19, 2001 — Observation Length: 216
Last Modified: 7:36 Friday, October 19, 2001 — Deleted Observations: 0
Protection: — Compressed: NO
Data Set Type: — Sorted: NO
Label:

-----Alphabetic List of Variables and Attributes-----

#	Variable	Type	Len	Pos	Format	Informat	Label
6	author	Char	50	115			First Author
8	cost	Num	8	8	DOLLAR9.2		Wholesale Cost
4	datesold	Num	4	32	MMDDYY8.	MMDDYY8.	Date Book Sold
9	listpric	Num	8	16	DOLLAR9.2		List Price
7	publishr	Char	50	165			Publisher
2	saleid	Num	8	0	8.		Sale ID
3	saleinit	Char	3	62			Sales Person Initial
10	salepric	Num	8	24	DOLLAR9.2		Sale Price
1	section	Char	26	36			Section
5	title	Char	50	65			

Macro variables

- `%let` keyword defines a macro variable and assigns it a value
- use & before macro variable name when referencing variable
- use `%eval` keyword to convert a macro variable's value to numeric
- when referencing macro variables in character literals, use double quotes

Macro variables example

```
%let repmonth=4;
%let repyear=2001;
%let repmword=%sysfunc(mdy(&repmonth,1,&repyear),monname9.)

data month&repmonth;
  set books.ytdsales;
  mosale=month(datesold);
  label mosale='Month of Sale';
run;

proc tabulate data=month&repmonth;
  title "Sales During &repmword &repyear";
  where mosale=&repmonth and year(datesold)=&repyear;
  class section;
  var salepric listpric cost;
  tables section all='**TOTAL**',
         (salepric listpric cost)*(n*f=4. sum*f=dollar9.2);
run;
* proc gchart data=month&repmonth ;
proc chart data=month&repmonth
           (where=(mosale < %eval(&repmonth+1) and
                  year(datesold)=&repyear));
  title "Sales Through &repmword &repyear";
  pie section / sumvar=salepric noheading ;
run;
```

Output

Sales During April 2001

	Sale Price		List Price		Wholesale Cost	
	N	Sum	N	Sum	N	Sum
Section						
Internet	145	$4,579.71	145	$4,680.75	145	$3,318.77
Networks and Communication	55	$1,633.01	55	$1,665.25	55	$1,177.46
Operating Systems	132	$4,016.45	132	$4,108.40	132	$2,916.03
Programming Languages	60	$1,835.07	60	$1,878.00	60	$1,330.98
Web Design	131	$4,015.50	131	$4,114.45	131	$2,910.87
TOTAL	523	$16079.73	523	$16446.85	523	$11654.09