

Introduction to SAS

Ely Kaizar

132M Baker Hall

ekaizar@stat.cmu.edu

July 15, 2002

Communicating with SAS

- Interactive
- Batch

SAS Structural Overview

- Types of Steps
 - Data Steps
 - Procedure Steps
 - Options/Settings Steps
- Syntax
 - All lines end with a semicolon
 - Use */*comment*/* to make comments

Data Step Syntax

1. Name your data set
2. Get data
3. Modify data set

Name Your Data Set

DATA *libname.dataname*;

Libraries:

- Temporary - library named WORK, default library
- Permanent - library named by you
LIBNAME *mylib 'file path/name'*;
- Naming Convention - library.dataname
NOTE: If you do not specify a library, then by default the library is WORK.

Get Data

- From a tab or space delimited file
INFILE '*file path/name*';
INPUT *numvar charvar \$...*;
(note the \$ after a non-numeric variable)
- From another sort of file
INFILE '*file path & name*';
INPUT *see help menu*;
- From previously created SAS data sets
SET *dataset1 [dataset2]*;
-or-
MERGE *dataset1 dataset2*;
BY *idnumber*;
- Type directly
CARDS *see help menu*;

Modify Data Set

1. Labeling and Formatting

- Labeling

```
LABEL myvar=" mylabel";
```

- Formatting

Before data step...

```
PROC FORMAT;
```

```
VALUE myformat 0=" label0" 1=" label1";
```

```
RUN;
```

During data step...

```
FORMAT myvar myformat;
```

Modify Data Set, cont

2. Selecting Only Parts of a Data Set

- Saving and Deleting Entire Variables

```
KEEP myvar1 myvar2;
```

```
DROP myvar1 myvar2;
```

- Saving and Deleting Observations

```
IF myvar=1;
```

```
IF myvar=0 then DELETE;
```

Modify Data Set, cont

3. Creating New Variables

- Creating Constants

myvar=1;

- Combining Other Variables

myvar3=*myvar1*+*myvar2*;

- if...then

IF *myvar1* < 10 then *myvar2* = 1;

ELSE IF *myvar1* < 20 then *myvar2* = 2;

ELSE *myvar2* = 3;

IF *myvar1* = . then *myvar2* = .;

Proc Step Syntax

```
PROC PROCNAME [DATA=mydata];
```

```
    OPTIONS;
```

```
RUN;
```

Common Proc Steps

- PROC PRINT;
- PROC SORT;
- PROC UNIVARIATE;
Numerical summary, Histogram
- PROC BOXPLOT;
- PROC GPLOT;
- PROC FREQ;
1- or 2-way tabulation, Test of proportions, χ^2
- PROC MEANS;
- PROC CORR;
Correlation
- PROC REG;
Regression
- PROC LOGISTIC;
Logistic Regression

Options/Settings

- LIBNAME
- OPTIONS NOCENTER
- OPTIONS LINESIZE PAGESIZE
- TITLE, TITLE2, ...