BIOS 145 **Lab #1** **1/14/98**

Introduction to SAS

TI: Ali Faiz
email: afaiz@bios.unc.edu

1. Web site up -- http://www.bios.unc.edu/classes/145

   - Possible URL change

   - Check regularly, especially “Course News”

2. Office hrs. TBA

3. Lab

   - 1st hour usually sample problems from HWs

   - 2nd hour for questions

   - No need to be here... lab is <u>optional</u>

   - Overheads will be available at HSL and downloadable from web page

   - Will try to with Edwards’ notes

BIOS 145 **Lab #1** **1/14/98**

- SAS background?

- What is SAS?

  -- Statistical Analysis System, latest v6.12

  -- “An integrated collection of data management, analysis, and reporting tools”

  -- Versions for a variety of environments, from mainframes to UNIX workstations to PC OS's like Windows 3.1, 95, NT, OS/2

- Need to know how to:

  -- Invoke the SAS system at your site: SPH Lab, HSL, etc.

  -- Manage files on your operating system

  -- Use the WWW

- UNC services/support

  -- 1 yr. license $60 from RAM Shop for students, faculty, staff, etc. Latest version on CD-ROM ~$60

  -- ATN Help Desk provides SAS help (962-5261, http://help.unc.edu)

  -- IRSS, ATN courses

  -- BIOS 111

- Two fundamental units of SAS programming:

  1. DATA step

  2. PROCs (procedures)

- DATA step

  -- Create data or read in data from a file

  -- Perform calculations

  -- Other data manipulations

  -- Create new datasets

- PROCs (procedures)

  -- Perform statistical, graphics, presentation, and some data management tasks

  -- Very useful for BIOS 145

- Types of data usable by SAS

  1. <u>Raw</u> <u>data</u>

     -- Plain text (ASCII) file with data entered by a text editor, word processor, etc.

     -- Only DATA steps can read and manipulate these

     -- Instructions in DATA step tell how to read in raw data and what data manipulations to perform, then convert the raw data to a SAS dataset

     -- SAS can also read in data from other formats, e.g. SPSS, dBASE, Lotus 123, Excel, etc. directly or via DBMSCOPY

  2. <u>SAS</u> <u>dataset</u>

     -- Contains not only the data, but instructions about how to display, what they mean, how created

     -- DATA steps and PROCs both can use

BIOS 145 **Lab #1** **1/14/98**

- How to run SAS?

  Two common ways are:

  1. Batch mode

     -- submits program to SAS (noninteractively) which processes it and returns results to files;

     e.g. typing **sas myprog.sas** at a command prompt submits the program myprog.sas

  2. Interactive mode

     -- Entered when SAS icon double-clicked

     -- See Program Editor, Log and Output Windows

     -- Used when submiting a SAS program from the Program Editor; results seen as they are produced rather than saved in files

     -- Can be slower than batch but more user-friendly

- What output produced?

  -- SAS programs can produce two types of output:

1. SAS Log: contains listing of the SAS program, any error messages/warnings, info on datasets used, etc. Always check after running

2. Procedure output: gives results of PROCs

-- Using myprog.sas example above, in batch mode, the Log is (typically) saved as **myprog.log** and the output as **myprog.lst**.

-- In interactive mode, Log is viewed in Log Window and output in Output Window

♦ A sample SAS program

-- A ";" denotes the end of a line of code

-- libname, options, cards statements

♦ Some data manipulations

-- Arithmetic Operators:

Operator	Operation	Examples
+	Addition	x+y
-	Subtraction	x-y
/	Division	rate/10
*	Multiplication	rate*10
**	Exponentiation	x**4

Can be used to define new variables: e.g,

z = (x - u) / s ;

-- Comparison operators: =, >, <, >=, <=, ^= (unequal)

e.g. if x<y

Can be used to obtain a subset of a dataset: if/where statements; e.g,

where (x<250);

would restrict the dataset to only those subjects with x<250. Here, "if" and "where" are interchangeable.

-- If/then statements

- Some useful PROCs

-- PROC PRINT, PROC CONTENTS

1. <u>Print</u> - prints some or all of a dataset

2. <u>Contents</u> - shows some characteristics of a dataset such as variables and # observations

e.g.,

```
proc print data=one;
   var a b c;
run;
```

```
proc contents data=one;
run;
```

-- PROC SORT

Sorts a dataset by specified variable(s)

e.g.,

```
proc sort data=one;
  var a;
run;
```

-- PROC MEANS, UNIVARIATE, FREQ

Gives various statistical results from a dataset

e.g.,

```
proc means data=one n nmiss mean std;
  var a b c;
run;

proc freq data=one;
  tables a b c a*b a*c / nocum missing missprint;
run;
```

(class/by statements)

BIOS 145 Lab #1 1/14/98

ATN Computer Classes

How do I register?
Call 962-1160.

Intro to SAS Programming
Dave Scocca teaches....(ask Dave)
Learn to build programs on PC or Statapps.
Tue, Jan-28, 3pm-5pm, #1L214

IRSS Computer Classes

SAS for Windows/95/NT and Unix: Getting Started in SAS (Part 1)
Tues, January 20
Jose Sandoval, IRSS
15 Manning Hall, UNC-CH, 3:30-5:00 pm
Info: Jose Miguel Sandoval, 919-962-0509, uirjms@frosty.irss.unc.edu

SAS for Windows/95/NT and Unix: Getting Started in SAS (Part 2)
Thurs, January 22
Jose Sandoval, IRSS
15 Manning Hall, UNC-CH, 3:30-5:00 pm
Info: Jose Miguel Sandoval, 919-962-0509, uirjms@frosty.irss.unc.edu

SAS for Windows/95/NT and Unix: Advanced SAS (Part 1)
Tues, January 27
Jose Sandoval, IRSS
15 Manning Hall, UNC-CH, 3:30-5:00 pm
Info: Jose Miguel Sandoval, 919-962-0509, uirjms@frosty.irss.unc.edu

SAS for Windows/95/NT and Unix: Advanced SAS (Part 2)
Thurs, January 29

Jose Sandoval, IRSS
15 Manning Hall, UNC-CH, 3:30-5:00 pm
Info: Jose Miguel Sandoval, 919-962-0509, uirjms@frosty.irss.unc.edu

Data Analysis Using SAS Interactively
Fri, January 30
Jose Sandoval, IRSS
15 Manning Hall, UNC-CH, 3:00-4:00 pm
Info: Jose Miguel Sandoval, 919-962-0509, uirjms@frosty.irss.unc.edu

♦ References:

DiIorio, Frank C. *SAS Application Programming: A Gentle Introduction*. Duxbury Press (PWS-Kent), Belmont, CA: 1991.

SAS Institute, Inc. *SAS Language: Reference, Version 6, First Edition*, Cary, NC: SAS Institute Inc., 1990.

SAS Institute, Inc. SAS Procedures Guide, Version 6, Third Edition, Cary, NC: SAS Institute Inc., 1990.