

## PSYC 5741/5751 Lab

### **Teaching Assistants:**

Roselinde Kaiser

kaiserr@colorado.edu

Office: Muenzinger D316C

Office Hours: Monday, 4 – 5

Allegra Smith

allegra.smith@colorado.edu

Office: Muenzinger E338

Office Hours: Tuesday, 1:30 – 3:30

### **Class Grade:**

- Average z-score from homeworks (drop 2 lowest) = 20%
- z-score for each test (2 mid-terms, 1 final) = 20% and 40%, respectively

### **Homeworks:**

- Please put your name on the back of the last page only
- We give partial credit and do not take off more than once if a mistake affects a series of calculations so it benefits you greatly to include your computations
- Remember to be explicit in answers and interpretations of analyses - We will not assume you have knowledge that is not demonstrated on the page
- Always attach your SAS code and output
- Homeworks will (theoretically) be posted on Wednesday. We will work on homeworks during lab on Thursdays and completed homeworks will be due at the beginning of class the following Tuesday

### **General Lab Format:**

- Return last homework and review common mistakes
- Brief presentation
- Go through new assignment to clarify what is being asked and what we'll be looking for when grading. Also go over SAS syntax when necessary.
- Note: Always bring your BuffOne card to lab so you can print out homework!

## Homework and SAS Basics

- Accessing your homework, past exams and the textbook:
  - These files are available through the web at [psych.colorado.edu/~gradstat](http://psych.colorado.edu/~gradstat). The username is “**model**” and the password is “**error**”. In addition to homework, old exams, and the text, you can download and print the PowerPoint presentations from class, under the folder “ChapterNotesFiles”.
  
- Getting Set Up:
  - Get on the mailing list:
    - Go to [psych.colorado.edu/mailman/listinfo/gradstat](http://psych.colorado.edu/mailman/listinfo/gradstat) and follow the instructions
  - Get a samiam account:
    - go to <http://cuconnect.colorado.edu>
    - enter CU Login Name and IdentiKey Password
    - select the “Courses” tab from along the top
    - on the left panel should be a ‘channel’ labeled “Plus”; select this
    - click "E-mail configuration" on the left of the screen
    - scroll down and select "Specialty Accounts"
    - validate for accounts by entering CUID and PIN
    - click "Make account on samiam"
      - After accepting User Responsibilities, choose a password, and the account will usually be made in the next 5 minutes.
  - Login to samiam by opening the X11 program and typing the command line:
    - `ssh -X [username]@samiam.colorado.edu`
  - After you hit enter, you’ll be asked to enter your samiam password. If it asks if you want to connect, type “yes” and hit enter.
  - **You only need to do this next part once:** type “`cp ~psycstat/autoexec.sas .`” at the prompt and hit enter. Don’t forget the space dot at the end. This will copy a file to your account that will run automatically each time you use SAS.
  
- From the Seuss Lab, you can open SAS by clicking on the “SASConnection” icon that appears on the dock. You’ll then be asked to enter your samiam username and password.
  - A second method to access SAS is to login to samiam, type “sas&” at the prompt, and SAS should launch. This second method makes printing from SAS a little easier.
  - The program editor, explorer and log window should be visible. The output window will be brought forward when a program is run from the editor window.
  - To change to a different window that you can see part of, just click the desired window. If the window you want is either hidden or was mistakenly closed, go to "View" in any window and select the window you want to see.
  
- From the CLIPR Macs, you can open SAS by opening the X11 program, logging on to samiam and typing “sas&” at the prompt.
  - To access your files that you saved on your samiam account, open the Fugu program and connect to samiam by entering “samiam” in the Connect To: box and your username in the username box. It will then prompt you for your password. Your samiam account will pop up, and you can drag and drop any of the files to the desktop. To print out your code, output or log files, open the files with the TextWrangler program. Save space/paper by printing two pages to a side.
  - You can also email your files to yourself and print them out elsewhere any time for free. Printing from the Seuss Lab costs money, but you do have a \$10 credit for copies on your BuffOne card.



## INTRO TO SAS PROGRAMMING

- Basic info
  - SAS is the statistical software package we will use for Grad Stat
  - We are using the old Unix version, as they stopped making updates for SAS on Macs. You can get newer, fancier versions for PC.
- SAS structure
  - The Windows
    - Editor window: where you write your code
    - Log window: where code you have run shows up, with any comments or errors
    - Explorer window: gives you access to your “libraries” (i.e. folders on your computer where datasets are stored)
    - Results window: skip back/forward to other analyses
    - Output window: where the results of your analyses show up
  - Editing
    - The text editor in this version of SAS is quite basic. For example, you cannot highlight selections and delete them as you would in most word processing programs. Instead, you have to depend on line commands that are entered into the blue line numbers on the left-hand side of the editor window.
    - **i** - inserts a line after the line you are on
      - **ib** - inserts a line before the line you are on
      - **i5** - inserts five lines after the current line; **ib5** inserts five lines before
    - **c** - copy a line; to paste, go to another line # and enter either **a** or **b** and then hit "return" to paste after or before that line, respectively; note this command only copies the line, it does not automatically delete it
    - to copy multiple lines, put **cc** on the first and last lines of the section to be copied; go to another line and enter **a** or **b** and "return" to paste
    - **d** or **dd** can be used to delete one line or multiple lines, respectively
    - **m** or **mm** can be used to move one or more lines to a different location; like copying, you have to choose another line and enter **a** or **b** and "return" to actually move the selection
  - Aside comments in programs. Before we embark on programming, you should know that you can place notes in a program that SAS will not read. Such notes can be very helpful to you later when you go back to try to figure out what you did or why you did it!
    - If you want SAS to ignore just one line, place one or two asterisks at the beginning of the line. Make sure to have a semicolon at the end of the line that way SAS knows where to stop ignoring text.
      - Example: **\*\* analysis of new data using 2 predictors;**
    - If you want SAS to ignore multiple lines, place " /\* " at the beginning and the reverse, " \*/ ", at the end of the selection you want ignored. Note: do not nest one set of these commands within another as SAS will get confused and ignore only the part between the first /\* and the first \*/.
      - Example of proper use: **/\* For this analysis, first test for the effects of variables while just controlling for the others. Then try the interactive model. Remember to mean deviate before running the interaction. \*/**
  - Basic SAS commands
    - LIBNAME: the name you give the folder on your account where your datasets are stored

- DATA: A data step is a command that allows you to create or transform variables in either a new or existing dataset
- PROC: A procedure command generates descriptive or inferential statistics as well as various other tasks (sorting, printing, etc)
- Creating and Using Datasets in SAS
  - There are 2 types of datasets that we can use - the # of tigers to the dataset name determines the type of dataset.
    - Temporary datasets have a one-tier name and exist only during the current SAS session. Example: **data=mydata**
    - Permanent datasets have a two-tier name. The first tier is the library the dataset resides in or is to be saved to (libname) and the second tier is the dataset's name. Example: **data=stat.auto**
  - Creating datasets in SAS
    - SAS datasets have the file extension .sas7bdat
    - To get a dataset in this format you can:
      - Enter data in manually
      - Input a .txt or .dat file
      - Read in a portable file from SPSS
      - Use a dataset already in that format

#### EXAMPLES

```
libname stats 'C:\My Documents\Stats';
stats.data1;
input id age cal;
datalines; /*This indicates that the data will follow*/
  1 7 1000
  2 10 1200
  3 15 1500
  4 20 1550
  5 22 1600
  6 25 1700
  7 30 1800
  8 40 1900
  9 50 2200
 10 60 1800
; /*However, for big datasets, this doesn't make much sense*/

data stats.data2;
  infile "C:\My Documents\Stats\example2.txt";
  input id age cal sex $; run;
```

```
/*I know, you really want to read in an SPSS file! Here is the easiest way to do that: From SPSS, save your dataset as a .por file, then do this:*/
libname base SPSS 'C:\Documents and Settings\Stats\mydata.por';
data stats.mydata;
set base._first_;
run;
```

- Transforming datasets in SAS

- We can add or remove new variables from a dataset in order to create a new dataset. You first specify the new dataset you want to create with a data step, then specify which dataset you want to “call” – that is, where the old data resides.

Example: the following calls the permanent dataset stat.auto into the step, transforms a particular variable (the transformations I used were random -- normally we'll have intelligent questions and transformations to apply), and then saves the contents of the old dataset plus the new variable into a temporary data file.

```

data stats.data3;
    set stats.data2;
    gender = 1*(sex = 'female')-1*(sex = 'male'); /*recoding a variable*/
    int = age*gender; /*Creating an interaction term*/
    months = age*12; /*Creates a variable with age in months rather than years*/
    child = 1*(age LT 20)-1*(age GE 20); /*Classify as either a child or not*/
    label child = 'Under age 20 or not';
    drop id; /*This tells SAS to drop the id variable from the new dataset*/
    keep age sex gender calcs int child; /*This tells SAS to only keep these variables in
    the new dataset*/run; /

proc print data = stats.data3 (obs = 5);
run; /*This will print all the variables for just the first 5 observations*/

```

- Statistics in SAS
  - For descriptive statistics for continuous variables in SAS, the primary commands to use are “proc univariate” or “proc means”. This will give us statistics such as mean, standard deviation and sample size, as well as things such as quantiles, skewness, and kurtosis.

```

proc univariate data = stats.data3; var age calcs; run; /*This is kind of overkill*/

proc means; /*This gives you N, mean, std dev, minimum and maximum
only*/var age; run;

proc means mean n var uss css clm;
var calcs;
Title 'Descriptive statistics for calorie intake'; run;

proc means mean n t prt; /*Will also do single sample t-tests, where null
hypothesis is mean = 0*/
class child; /*will give you means for both child and adult*/
var calcs;
run;

```

- For categorical variables, the primary command is “proc freq”, which gives you frequencies

```

proc freq data = stats.data3;
table sex/chisq;
run;
proc freq data = stats.data3;
tables sex*child /chisq; /*This gives you crosstabs*/

```

**run;**

- Regression analyses

```
proc reg data = stats.data3;  
model calcs = age; /*Age is the independent variable, calcs is the dependent  
variable*/run;
```

- Sorting dataset by a variable(s)

- Sometimes you will want to run a procedure separately for different levels of a variable. To do this you must first sort the dataset by the variable. Then run the procedure command also sorted by the variable.

```
proc sort data=<dataset name> out=<new name>;  
by zzz;  
proc means data=<new name>;  
by zzz;  
var yyy xxx;  
title 'MEANS FOR YYY, XXX SORTED BY ZZZ';  
run;
```

- Graphing in SAS

```
proc gchart;  
vbar age calcs /levels = 4;  
Title 'Vertical Bar Charts with 4 levels';  
run;  
proc gplot data = stats.data3;  
plot calcs*age; /*the first variable is on the y axis, the second on the x-axis*/Title  
'Scatterplot of Age by Calories';  
run;  
proc gplot;  
plot calcs*age;  
symbol1 v = circle c = blue i = r;  
Title 'Scatterplot with regression line';  
run;
```

- Submitting and recalling programs
  - To run a program, hit F3 or go to "Run" and then "Submit". To recall a program, hit F4 or "Run" and "Recall last submit".

➤ Saving and printing programs and output. You should always save your programs that way you can just edit them later if you find mistakes rather than having to retype the whole thing. Plus, in order to print your work, you will have to save the file and then print from the samiam directory.

- Saving programs. In the editor window, go to "File" then "Save As". A dialog box should appear. On the left-hand side is the directory that SAS wants to save your program in -- this directory will be whatever directory you were in on samiam before you opened up SAS. If you are in the wrong place, you can type a tilda (~) then your login name to get you back to your home directory (note: don't put a space between the tilda and your login). Type in the name you want for the file and always use the extension ".sas" for a program! I suggest making one file per homework (not one file per problem).

- Saving output. Basically the same procedure as above except that you are doing it from the output window instead of the editor window and the extension you should use is ".lst". (Make sure you change the extension to .lst or you will overwrite your program!).
- To print go back to the samiam window (in the X11 program) and type the following at the prompt:

**print <program filename.sas> <output filename.lst> then hit "return"**

This command will print 4 pages to each sheet of paper and will print the program and the output files back to back. (I always want your program file in addition to your output for homeworks so this is a good way to remember both.)

Canceling a print job. Type "**lpq**" and "return" at the samiam prompt to list the jobs waiting in the printing queue. You should see your login name somewhere along with a job number for your print job. (After looking over the list you may have to hit "return" again to get a samiam prompt back). To remove a job, type "**lprm <your job #>**" and "return".

- Quit SAS and samiam by hitting "File" then "Exit" in the editor window. Then quit the X11 and/or Terminal programs.
- Log off the computer by hitting the "logout" icon on the desktop (if you accidentally logout while SAS or SPSS is still open, notify me or the ITS person in E312 immediately so the resulting runaway process can be killed).
- Some Good SAS resources:
  - The Little SAS Book: A Primer, 3<sup>rd</sup> Edition. By Lora D. Delwiche & Susan J. Slaughter
  - <http://www.ats.ucla.edu/stat/sas/modules/>
  - <http://support.sas.com/onlinedoc/913/docMainpage.jsp>