Statistics 157: Lab #6
Winter 2018; 10 pts

NAME: Sarah Ruckman

ID: (last 4 digits only) 7194

INSTRUCTIONS:

- Save your SAS program file as *lab6w18.sas*.
- You will need the following data files (They are available for download on *iLearn* under Data Files.):

    sem1.dat   sem2.dat   sem3.dat   sem4.dat   sem5.dat   sem6.dat

- When you have completed the lab, please be sure you straighten your work area, and sign the lab roster sheet.

This lab is designed to give the students practice with writing some basic SAS macros.

Please go slowly, reading everything carefully and making sure you understand each step before moving on to the next one! You are excellent students and can do this!!

A Suggestion: When using SAS macros, it is difficult to keep track of all the cross-referencing. My suggestion is to first print a hard copy of your program. Then using multiple highlighters, color-code the program. For example: Suppose you are passing across the name of an existing SAS dataset using the reference **olddata**. Go through your program and highlight (in the same color) everywhere **olddata** is used.

NOTE: You do not have to type in the comments for this lab! (You might want to at some point in time for future reference!)

NOTE: Comments within a SAS macro are enclosed by

    %*   Commment here ;

Suppose you work in the registrar's office of a university. You have decided that you want to write a SAS program to make data entry of transcript information and grade computation more efficient. You have decided to write a series of SAS macros to achieve your goal. It is assumed that the data for 6 semesters for a particular student has been entered into 6 data files: *sem1.dat, sem2.dat, sem3.dat, sem4.dat, sem5.dat* and *sem6.dat*. (These files are available for download on Blackboard under **Data Files**. Make sure you have them before proceeding.)

NOTE: We will add some new options to generate diagnostics during the execution of the SAS macros. These options are *mtrace, mlogic* and *mprint*.

PLEASE clear your log file before every submission of your new code! (To do this, click to activate the log window and then select *Edit → Clear all*. Be sure you save your file before each execution. You may have to exit SAS and clear it's internal footprints more frequently than before.

NOTE: The infile statement below is

    infile "c:\linda\winter2018\w18157\datafiles\sem&i..dat" firstobs=4;

1

Be sure to change the path to the location where you have saved your data files!

- The &i allows the value of i to change according to the do loop index variable value.

- You do need the 2 dots (not a typo).
  ♣ The first dot indicates the end of the parameter name (in this case i) and tells SAS to execute the change in the value of i according to the do loop index variable value. (For example, you might have a do loop index variable named jj. Then you would need sem&jj..dat so SAS would know the parameter is jj not just j.)
  ♣ The second dot is part of the extension .dat.

1. The first thing you want to accomplish is the reading in of the raw data. For that, create a SAS macro called *indata1* as follows: (You do not need to enter the comments at this time, but be sure you READ and understand them!)

```
options ls=78 nocenter nodate ps=55 mprint mtrace nodate nonumber formdlim = '#';
/*
************************************************************
************************************************************
**   Macro indata1    macro to read individual data files into    **
**                     individual SAS dataset and a cumulative     **
**                     SAS dataset. The data files contain typical **
**                     college transcript information (including   **
**                     course department, number, name, number of  **
**                     semester units and final grade). Once the   **
**                     data has been read in, the macro computes   **
**                     gpa for current semester, cumulative gpa    **
**                     for all courses, cumulative gpa for courses **
**                     in the student's major, cumulative gpa for  **
**                     courses numbered below 200 and cumulative   **
**                     gpa for courses numbered above 200.         **
************************************************************
************************************************************
**   Variable specification:                                       **
**      olddata    name of SAS dataset to be used for computing    **
**                 cumulative gpa for courses numbered below 200   **
**                 and cumulative gpa for courses numbered above   **
**                 200                                             **
**      newdata    name of SAS dataset to be used for computing    **
**                 cumulative gpa for all courses                  **
**      semnum     name of SAS dataset to be used for computing    **
**                 gpa for all courses in current semester         **
**      number     number of semesters of data                    **
**      name1      name of student                                 **
************************************************************
************************************************************  */
%macro indata1(olddata,newdata,semnum,number,name1);
%* Read in and print data semester by semester;
%* Create a new temporary SAS data set for each semester;
  %do i = 1 %to &number;
```

```
        data &semnum&i;
%* To enter a data file, you need 2 periods as shown below;
%* Be sure to change the path to your data files;
    infile "c:\linda\winter2018\w18157\datafiles\sem&i..dat" firstobs=4;
      input dept $ 1-4 crsnum crsdesc $ 16-40 units grade $;
proc print noobs;
    var dept crsnum crsdesc units grade;
%end;
%mend indata1;
```

Now let's try to invoke your macro.

NOTE: In the following, the names *restr*, *cumul* and *semstr* are just dummy SAS dataset names.

(i) First, try reading in and printing out the first data set. Add the following lines of code to your program (after the %mend indata1 statement) and then execute your program.

```
/*
*****************************************************************
*****************************************************************
**      Test the macros using the six semester data files      **
*****************************************************************
*****************************************************************
*/
/* indata1(olddata,newdata,semnum,number,name1);
   Acording to the macro call below, the parameters will take on
   the following names:
      olddata   will take on the name    restr
      newdata   will take on the name    cumul
      semnum    will take on the name    semstr
      number    will take on the value   1
      name1     will take on Your name (put in your actual name
                where I have written "Your name")
*/

%indata1(restr,cumul,semstr,1,Your name);
run;
quit;
```

Your output should look like:

dept	crsnum	crsdesc	units	grade
MAST	8	Calculus I	5	C
ENGL	5	English Composition I	3	A
CSCI	25	Intro to CS	4	A
PHYS	8	Physics I	5	B

Make sure your output matches the above before going to the next step! Have your neighbor or TA initial here _____

(ii) Now test your macro to see if it will read in all six data sets. Change the macro execution (invocation) statement to: (Change the 1 to a 6!)

3

`%indata1(restr,cumul,semstr,6,Your name);`

and re-execute your program.

When you get your new output, have your neighbor or TA initial here ____DA____

2. Now that you have successfully read in your data, its time to make headings for our transcript. Create a new macro, called *semhead*. Place it in your program file before the macro *indata1*. Notice that in a macro, the titles are in double quotes (as opposed to the usual single quotes). This allows SAS to actually execute them as executable statements (not just print them as verbatim statements) and change that values of the parameters within them.

```
/*
*******************************************************************
*******************************************************************
**   Macro semhead   macro to print headings for each semester   **
**                                                               **
**   Variable Specification                                      **
**      semnum        number of current semester                 **
*******************************************************************
*******************************************************************
*/

%macro semhead(semnum,name1);
   title1 "Transcript for &name1, Semester &semnum";
   title2 "Wishful University";
   title3 "Anywhere, Anystate";
%mend semhead;
```

Now modify your macro *indata1* to include heading information. This will simply be the addition of a single line of code (plus a comment line) in the macro.

```
%macro indata1(olddata,newdata,semnum,number,name1);
%* Read in and print data semester by semester;
%* Use do loops to create a new temporary SAS data set for each semester;
   %do i = 1 %to &number;
      data &semnum&i;
%* *********************************************;
%* NEW LINES OF CODE FOLLOW                    ;
%* Create semester headings;
      %semhead(&i,&name1);
%* To enter a data file, you need 2 periods as shown below;
%* Be sure to change the path to your data files;
      infile "c:\linda\winter2018\w18157\datafiles\sem&i..dat" firstobs=4;
      input dept $ 1-4 crsnum crsdesc $ 16-40 units grade $;
%* Print the data as a check;
proc print noobs;
   var dept crsnum crsdesc units grade;
%* Close the loop;
%end;
%mend indata1;
```

Save your file and re-submit to see your new results. (You do not need to change the executable *indata1* statement!

4

Sample Output for Semester 1:

```
Transcript for Joe Student, Semester 1
Wishful University
Anywhere, Anystate
Individual Semester Information
dept   crsnum   crsdesc                 units   grade
MAST     8      Calculus I                5       C
ENGL     5      English Composition I     3       A
CSCI    25      Intro to CS               4       A
PHYS     8      Physics I                 5       B
```

Make sure your output matches the above before going to the next step! Have your neighbor or TA initial here ____DA____

3. OK, now its time to code letter grades into numerical values and compute total points for each course. (This information will be needed to compute grade point averages!) Create a new macro *gradcode* and place it in your program AFTER the macro *semhead*. (NOTE: You can use a regular if-then-else structure within the macro!)

```
/*
*******************************************************************
*******************************************************************
**   Macro gradcode   macro to code letter grades to numbers   **
*******************************************************************
*******************************************************************
*/

%macro gradcode;
        if grade = "A" then pts = 4;
   else if grade = "B" then pts = 3;
   else if grade = "C" then pts = 2;
   else if grade = "D" then pts = 1;
   else pts = 0;
%mend gradcode;
```

Now incorporate this into your macro *indata1* as follows:

```
%macro indata1(olddata,newdata,semnum,number,name1);
%* Read in and print data semester by semester;
%* Create a new temporary SAS data set for each semester;
   %do i = 1 %to &number;
        data &semnum&i;
%* Create semester headings;
        %semhead(&i,&name1);
%* To enter a data file, you need 2 periods as shown below;
%* Be sure to change the path to your data files;
        infile "c:\linda\spring2017\s17157\datafiles\sem&i..dat" firstobs=3;
        input dept $ 1-4 crsnum crsdesc $ 16-39 units grade $;
%* *********************************************** ;
%* NEW CODE FOLLOWS                               ;
```

```
%* Transform letter grades to numeric values;
      %gradcode;
%* Calculate total points earned in a semester;
      tpts = units*pts;
      title4 "Individual Semester Information";
proc print noobs;
   var dept crsnum crsdesc units grade tpts;
%end;
%mend indata1;
```

Save and execute your program. Your output for Semester 6 should look like:

```
Transcript for Joe Student, Semester 6
Wishful University
Anywhere, Anystate
Individual Semester Information
dept      crsnum      crsdesc                   units     grade     TPTS
MAST      252         Advanced Calculus II        3         B         9
PetE      302         Design I                    4         A        16
CHEM      252         Organic Chemistry II        4         B        12
PetE      315         Special Topics              4         A        16
GERM      101         Intro to German             3         D         3
```

Check to make sure your output matches this output! Have your neighbor or TA initial here _____

Notice that with every addition you have made, the execution of the macro *indata1* has not changed. It has remained as

`%indata1(restr,cumul,semstr,6,Your name);`

NOTE: If you only wanted to use say the first 3 semesters of data, the only change that would be required is to change the 6 to a 3:

`%indata1(restr,cumul,semstr,3,Your name);`

Question: Could you modify this program select, say, semesters 1, 3 and 5??????

4. Now for the BIG moment - the computation of the individual semester gpa. Create a new macro called **gpa1**. Place this macro before the **indata1** macro!

NOTE: BE SURE YOU UNDERSTAND HOW THIS MACRO IS WORKING! REMEMBER, GPA = (Total # grade points earned)/(total # hours)

```
/*
***************************************************************
***************************************************************
** Macro gpa1   macro to compute sum of semester hours and sum **
**              of earned grade points in separate SAS datasets,**
**              merge the SAS datasets into one, generate the  **
**              number of the semester, and print semester,    **
**              total number of course hours, total number of  **
```

6

```
**                      earned grade points and gpa                **
**   Variable Specification:                                       **
**     whdata  name of existing SAS dataset to be used             **
**     i       number of semester                                  **
**     sum1    total number of course hours                        **
**     sum2    total number of earned grade points                 **
**     gpa     grade points average = sum1/sum2                     **
**     out1    name of SAS dataset where information is to be       **
**             stored                                              **
*******************************************************************
*******************************************************************
*/


%macro gpa1(whdata,i,sum1,sum2,gpa,out1);
%* Generate total number of COURSE HOURS;
proc means data = &whdata&i noprint;
   var units;
%* Output the sum of the course hours to a
   dataset named "a" and in a variable named
   sum1;
%* NOTE: the &i allows you to automatically generate a new "a" output
   file each time through the do loop. The first time through the
   output file will be a1; second one will be a2, etc. So you will
   have an "a" output file for each of the semesters ;

   output out= a&i sum = &sum1;
run;
%* Generate total number of EARNED GRADE POINTS;
proc means data = &whdata&i noprint;
   var tpts;
%* Output the sum of the course hours to a dataset named "b"and in a
   variable named sum2;
%* NOTE: the &i allows you to automatically generate a new "b" output
   file each time through the do loop. The first time through the
   output file will be b1; second one will be b2, etc. So you will
   have a "b" output file for each of the semesters ;

   output out = b&i sum = &sum2;
run;
%* Merge above datasets (This will put the total number of course
   hours and total number of earned grade points in the same file so
   you can do calculations with them. Also assign semester number
   and compute gpa. Print information;
%*.   a&i  temporary SAS data set contains total number of course hours for semester i
      b&i temporary SAS data setcontains totla number of grade points earned for semester i
      sum1  variable representing the total number of course hours
      sum2  varibale representing the total number of grade points
      gpa = (total number of grade points)/(total number of course hours) ;

data &out1&i;
   merge a&i b&i;
```

7

```
       semster = &i;
       &gpa = &sum2/&sum1;
proc print noobs;
   var semster &sum1 &sum2 &gpa;
run;
%mend gpa1;
```

Incorporate the gpa1 macro into your macro *indata1*:

```
%macro indata1(olddata,newdata,semnum,number,name1);
%* Read in and print data semester by semester;
%* Create a new temporary SAS data set for each semester;
   %do i = 1 %to &number;
       data &semnum&i;
       %semhead(&i,&name1);
       infile "a:sem&i..dat" firstobs=3;
       input dept $ 1-4 crsnum crsdesc $ 16-39 units grade $;
       %gradcode;
       tpts = units*pts;
       title4 "Individual Semester Information";
proc print noobs;
   var dept crsnum crsdesc units grade tpts;
%* Create cumulative grades file and individual semster gpa;

%*   ****** NEW STUFF STARTS HERE ********* ;
%* Compute individual gpa information;
%* Recall: sumsem_1 = sum of course points;
%*         sumsem_2 = sum of units;
%*         sem_gpa = sumsem_1/sumsem_2;
%*              = gpa for current semester;

%gpa1(&semnum,&i,sumsem_1,sumsem_2,sem_gpa,sd);

%* ******** NEW STUFF ENDS HERE ********** ;
%end;
%mend indata1;
%indata1(restr,cumul,semstr,6,Your name);
run;
quit;
```

Save and execute your program and complete the following for Semester 6:

semster	sumsem_1	sumsem2	sem_gpa
6	18	56	3.1111

You do **not** need to turn in this lab worksheet. Please sign the lab roster for his week to get credit for this lab. Please make sure your work area is neat and have a good weekend!

Linda & Luke

DATA FILES

```
Name this file sem1.dat
12345678901234567890123456790123456789012345678901234567890
Dept   Course#  Course Name            Units   Grade
MAST   8        Calculus I             5       C
ENGL   5        English Composition I  3       A
CSCI   25       Intro to CS            4       A
PHYS   8        Physics I              5       B


Name this file sem2.dat
12345678901234567890123456790123456789012345678901234567890
Dept   Course#  Course Name             Units   Grade
MAST   21       Calculus II             5       B
ENGL   10       English Composition II  3       A
PetE   10       Intro to Pet Engr       2       A
PHYS   25       Physics II              4       B
CHEM   8        Inorganic Chemistry     4       A


Name this file sem3.dat
12345678901234567890123456790123456789012345678901234567890
Dept   Course#  Course Name             Units   Grade
MAST   22       Calculus III            4       A
PetE   175      Statics                 4       C
CSCI   150      C+ Programming          4       A
CHEM   10       Inorganic Chemistry II  4       B


Name this file sem4.dat
12345678901234567890123456790123456789012345678901234567890
Dept   Course#  Course Name              Units   Grade
MAST   171      Linear Algebra           4       B
MAST   172      Differential Equations   3       B
PetE   150      Technical Writing I      2       A
PetE   175      Dynamics                 4       C
CSCI   151      C++ Programming          4       A


Name this file sem5.dat
12345678901234567890123456790123456789012345678901234567890
Dept   Course#  Course Name             Units   Grade
MAST   215      Engineering Statistics  3       A
MAST   251      Advanced Calculus I     3       B
PetE   151      Technical Writing I     2       A
CHEM   251      Organic Chemistry I     4       A
PetE   202      Heat Transfer           4       B


Name this file sem6.dat
12345678901234567890123456790123456789012345678901234567890
Dept   Course#  Course Name             Units   Grade
MAST   252      Advanced Calculus II    3       B
PetE   302      Design I                4       A
CHEM   252      Organic Chemistry II    4       B
PetE   315      Special Topics          4       A
GERM   101      Intro to German         3       D
```