Syllabus

Data management and SAS Programming Language

Epidemiology/Public Health 576D

Fall 2014

Time: Tuesday and Thursday 11:00 – 12:15 pm

Room: Drachman Hall, Room A319

Instructor: Angelika Gruessner, PhD
Professor Epidemiology & Biostatistics
A224 Drachman Hall
Telephone: 626-3118
acgruess@email.arizona.edu

Office hours: Monday, Wednesday 11.00-12.15pm (or by appointment)

Teaching Assistants: William John Degnan <wjed@email.arizona.edu>

TA office hours: TBA

Catalog Description: This course will introduce students to the fundamentals of data management and analysis using the SAS system. Emphasis will be placed on the management of large distributed data sets and data manipulation, including reading, processing, recoding, and reformatting of data. The approach will be to teach by example, with an emphasis on hands-on learning. The course will cover the specific topics listed in the course objectives below. Each participant should afterwards be able to take and pass the ‘Base’ and possibly also the ‘Advanced Programmer for SAS 9’ Certification offered through the SAS Institute. Therefore, topics of this course mirror those appearing on the certification test: advanced programming, using SQL with SAS, and optimizing SAS programs.

Course Prerequisites: Basic computer literacy, CPH576A or permission of the instructor
Course Learning Objectives: At the end of the course, you should be able to:

- Demonstrate critical data management and programming skills using the SAS system
- Pass the global Base SAS certification
- Demonstrate advanced programming techniques in SAS such as creating and using indexes, combining data horizontally and vertically, compressing SAS data sets, etc.
- Perform macro programming in SAS such as text substitution in code, automating and customizing the production of SAS code, conditionally or iteratively constructing SAS code, etc.
- Demonstrate skills in SQL with SAS and optimizing SAS programs such as querying and subsetting data, summarizing and presenting data, combining tables, including complex joins and merges
- Being able to use SAS, R, SPSS, Stata interchangeable

MPH Competencies Covered: At the end of the course, you should be able to:

Analytical Skills:

- Define a problem
- Determine appropriate uses and limitations of data
- Select and define variables relevant to defined public health problems
- Evaluate the integrity and comparability of data and identify gaps in data sources

Communication Skills:

- Communicate effectively both in writing and
- Interpret and present accurately and effectively demographic, statistical, and scientific information for professional and lay audiences adapting and translating public health concepts to individuals and communities

Basic Public Health Skills:

- Understand research methods in all basic public health sciences
- Apply the basic public health sciences including behavioral and social sciences, biostatistics, epidemiology, environmental public health, and prevention of chronic and infectious diseases and injuries

Data Management Competencies Covered:

- Identify appropriate tools to address specific scientific questions
- Demonstrate excellent presentation skills and the ability to explain statistical concepts and findings to a general scientific audience
- Demonstrate understanding of methods of data analysis and data monitoring
**Course Notes:** A webpage has been created for this class using the Desire 2 Learn (D2L) interface. The course website contains the syllabus, class notes, and all information you need for homework assignments.

Students will be given a copy of the power point presentation and handouts for the block. They may also be given additional examples of specific SAS programs which will also be discussed in class. Students are responsible for **ALL material** distributed during the semester.

To access the **576D** website, login at: [http://d2l.arizona.edu](http://d2l.arizona.edu)

- Click the ‘UA NetID’ Login
- Enter your NetID and password, the same way you would access your UA student account
- Under “My Courses” click on **CPH EPID 576D FA14 001**
  - News: This section contains any class announcements, such as changes in the homework schedule, etc.
  - Content: Access the syllabus, class notes, homework assignments and supplemental information in this section.

For further information on how to use the D2L interface, go to:

[http://www/help.D2l.arizona.edu](http://www/help.D2l.arizona.edu)

Note that if you do not have a UA NetID, please see me so that I can give you access to the D2L site.

**Recommended Texts:**

- Online documentation: [http://support.sas.com/documentation/index.html](http://support.sas.com/documentation/index.html)

**Course Requirements**

1. **Review the notes**
   a. Please bring questions about the notes with you to the lecture

2. **Homework**
   a. The homework assignment for each block is included in D2L.
   b. Students will be required to complete all assigned problems and to submit the results (SAS code) **by the assigned date** into the D2L drop box. Students are expected to submit code and successful runs, unless they indicate and explain their errors/problems in their program(s).
   c. Scoring: Each homework is worth 3 points. Partial credit will be given if an honest attempt at the problem was made even without the correct answer. No credit (zero points) is given if no attempt was made to solve the problem.
d. **Remember to put your name on the front page.**
e. Answers will be discussed in class.
f. Keep copies of all of your code and assignments. Those will help you for the final assignment.

3. **Examination**
   a. One midterm assignment will be given to you. It will be a take-home, project style assessment that involves knowledge, concepts, and skills cover in block 1-7.

4. **Final Project**
   a. The final project is meant to give you an opportunity to show that you understand the problems of data management
   b. Students will be given a problem which has to be solved outside class.
   c. Students are expected to complete these independently. Specific questions about these items can be raised during class time.

5. **Extra Credit**
   a. No extra credit is available

6. **USB key for all the material**

**Grading/Student Evaluation:** Homework, midterm and the final project contribute to your final grade as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>40%</td>
</tr>
<tr>
<td>Midterm</td>
<td>20%</td>
</tr>
<tr>
<td>Final Project</td>
<td>40%</td>
</tr>
</tbody>
</table>

Final grades are based on the following point system:

- A = 90-100%
- B = 80-89%
- C = 70-79%
- D = 60-69%
- E = 59% or less

The instructor reserves the right to revise this scale, if necessary.

**Class Attendance/Participation:** Class attendance is strongly encouraged. If a student misses class, they are responsible for meeting all course deadlines, and for working with other students, the TAs and the instructor (during office hours) to catch up. All holidays or special events observed by organized religions will be honored for those students who show affiliation with that particular religion. Absences pre-approved by the UA Dean of Students (or Dean's designee will be honored.)
**Course Schedule:** Any changes to the following schedule will be announced in lecture or the D2L site. You are responsible for obtaining information on any changes, even if you miss class.

<table>
<thead>
<tr>
<th>BLOCK</th>
<th>TOPIC</th>
<th>SUBTOPICS</th>
<th>HOMEWORK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction to SAS</td>
<td>Structure of SAS&lt;br&gt;DATA and PROC steps&lt;br&gt;SAS Windows Environment&lt;br&gt;Variable definitions&lt;br&gt;Syntax rules</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Introduction to simple data input</td>
<td>SAS: Program Data Vector, Syntax Rules&lt;br&gt;INFILE statement&lt;br(INPUT statement&lt;br&gt;LIBNAME statements</td>
<td>Problem 1</td>
</tr>
<tr>
<td>3</td>
<td>Data manipulation, conditional processing and missing values</td>
<td>Creating new variables and transformations&lt;br&gt;IF/THEN/ELSE</td>
<td>Problem 2</td>
</tr>
<tr>
<td>4</td>
<td>Simple statistical SAS procedures</td>
<td>MEANS, FREQ, UNIVARIATE, TTEST, CORR, REG</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>SAS ARRAY computations</td>
<td>ARRAY</td>
<td>Problem 3</td>
</tr>
<tr>
<td>6</td>
<td>SAS FUNCTIONS</td>
<td>Mathematical functions&lt;br&gt;Character functions&lt;br&gt;Pseudo random number generator</td>
<td>Problem 4</td>
</tr>
<tr>
<td>7</td>
<td>SAS dates</td>
<td>Definition and date and time functions</td>
<td>Problem 5</td>
</tr>
<tr>
<td>8</td>
<td>Restructuring data sets</td>
<td>LAG and RETAIN</td>
<td>Problem 6</td>
</tr>
<tr>
<td>9</td>
<td>SAS labels and formats</td>
<td>PROC FORMAT, format library</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SAS output and SAS graphics</td>
<td>SAS ODS (Output delivery system)&lt;br&gt;RTF, HTM L, PDF formats</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Data cleaning techniques</td>
<td></td>
<td>Problem 7</td>
</tr>
<tr>
<td>12</td>
<td>Sorting, merging and update of data sets</td>
<td>MERGE, UPDATE</td>
<td>Problem 8</td>
</tr>
<tr>
<td>14</td>
<td>SAS engines and interfaces</td>
<td>EXCEL, ACCESS, SPSS, STATA</td>
<td>Problem 9</td>
</tr>
<tr>
<td>15</td>
<td>Database management and manipulation, table lookup</td>
<td>PROC SQL</td>
<td>Problem 10</td>
</tr>
<tr>
<td>16</td>
<td>SAS macro variables and macro programming</td>
<td>Improving programming and introduction to simulations</td>
<td>Problem 11</td>
</tr>
<tr>
<td>17</td>
<td>Introduction to matrix language</td>
<td>SAS IML</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Final SAS project</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Statistical Software:** You will need to use SAS extensively

A set of modules for learning SAS are on the D2L site. All students enrolled in the course have access to these modules.

**Computer Labs:** SAS are available for public use at two locations:

- Drachman Hall Computing Lab: Drachman A319, open weekdays, from 8-5. There are printers available free of charge if you want to print homework, but you must supply your own paper.
• Arizona Health Sciences Library Computer Lab: AHSC 2150, open every day 6am-midnight. These computers are behind the information/reference desk on the main floor. The first couple of banks of machines is not part of the lab, but is rather used for lit searching, etc. The lab is the ‘walled off’ section of computers behind the first couple of banks. You may print output here for a fee. Note that these are public facilities, and may or may not be crowded on a given day. Students must register to use the AHSC Library Computer Lab at the Library Information Desk. A University of Arizona Catcard is required.

Purchasing SAS: SAS can be ordered from the UA BookStore. The cost is $95 per year. The link with the necessary information is [http://www.uofabookstores.com/uaz/CATS/License_SAS.asp](http://www.uofabookstores.com/uaz/CATS/License_SAS.asp).

SAS University Edition: This is a new version for Universities. It is free. I will give out directions how to set it up. It should be available for Windows, Macintosh, and Linux operation systems.

Tips for Succeeding in the Course:

1. Attend class and your discussion section
2. Read the lecture notes before coming to class
3. Ask questions about the notes and textbook in class
4. Do your homework early
5. Check your homework answers against the answer key
6. Turn in your work on time
7. Ask questions until you understand the material

Required Statements:

Communications: You are responsible for reading emails sent to your UA account from your professor and the announcements that are placed on the course web site. Information about readings, news events, your grades, assignments and other course related topics will be communicated to you with these electronic methods. The official policy can be found at: [http://www.registrar.arizona.edu/emailpolicy.htm](http://www.registrar.arizona.edu/emailpolicy.htm)

Disability Accommodation: If you anticipate issues related to the format or requirements of this course, please meet with me. I would like us to discuss ways to ensure your full participation in the course. If you determine that formal, disability-related accommodations are necessary, it is very important that you be registered with Disability Resources (621-3268; drc.arizona.edu) and notify me of your eligibility for reasonable accommodations. We can then plan how best to coordinate your accommodations. The official policy can be found at: [http://catalog.arizona.edu/2013%2D14/policies/disability.htm](http://catalog.arizona.edu/2013%2D14/policies/disability.htm)
**Academic Integrity:** All UA students are responsible for upholding the University of Arizona Code of Academic Integrity, available through the office of the Dean of Students and online: The official policy found at: [http://deanofstudents.arizona.edu/codeofacademicintegrity](http://deanofstudents.arizona.edu/codeofacademicintegrity)

**Classroom Behavior:** Students are expected to be respectful of the instructor, TAs and other students at all times (including limited talking, no reading newspapers, etc.). Cell phones should be in the mute or vibrate position. If you must take an emergency call during class please leave class quietly to speak with the caller (do not leave and return more than once as this disrupts the rest of the class). Please do not text during class. Students may use their laptops during class only for course related material.

The Dean of Students has set up expected standards for student behaviors and has defined and identified what is disruptive and threatening behavior. This information is available at: [http://deanofstudents.arizona.edu/disruptiveandthreateningstudentguidelines](http://deanofstudents.arizona.edu/disruptiveandthreateningstudentguidelines)

Students are expected to be familiar with the UA Policy on Disruptive and Threatening Student Behavior in an Instructional Setting found at: [http://policy.arizona.edu/disruptive-behavior-instructional](http://policy.arizona.edu/disruptive-behavior-instructional) and the Policy on Threatening Behavior by Students found at: [http://deanofstudents.arizona.edu/sites/deanofstudents.arizona.edu/files/Disruptive_threat_bklt_2012.pdf](http://deanofstudents.arizona.edu/sites/deanofstudents.arizona.edu/files/Disruptive_threat_bklt_2012.pdf)

**Grievance Policy:** Should a student feel he or she has been treated unfairly, there are a number of resources available. With few exceptions, students should first attempt to resolve difficulties informally by bringing those concerns directly to the person responsible for the action, or with the student's graduate advisor, Assistant Dean for Student and Alumni Affairs, department head, or the immediate supervisor of the person responsible for the action. If the problem cannot be resolved informally, the student may file a formal grievance using the Graduate College Grievance Policy found at: [http://grad.arizona.edu/academics/policies/academic-policies/grievance-policy](http://grad.arizona.edu/academics/policies/academic-policies/grievance-policy)

**Grade Appeal Policy:** [http://catalog.arizona.edu/2013-14/policies/gradappeal.htm](http://catalog.arizona.edu/2013-14/policies/gradappeal.htm)

**UA Smoking and Tobacco Policy:**

The University's "Smoking and Tobacco Policy" is designed to promote the health and wellness of all members of the University community, including visitors to campus, and it will prohibit the use of tobacco- and nicotine-containing products on property owned or controlled by the UA. This includes the main campus, the Arizona Health Sciences Center, the Phoenix Biomedical Campus, UA South, all satellite campuses, University vehicles, and any property leased by the UA. Smoking cessation aids, such as nicotine gum, patches, and nasal sprays, will be
permitted, but all other forms of tobacco or nicotine - including pipes, cigars, cigarettes and e-cigarettes, all types of smokeless tobacco, and water pipes - will be prohibited. The latest version of the policy is available at: http://policy.arizona.edu/sites/default/files/Tobacco-Free.pdf

**Syllabus Changes:** Information contained in the course syllabus, other than the grade and absence policies, may be subject to change with reasonable advance notice, as deemed appropriate.

**Plagiarism:** What counts as plagiarism?

- Copying and pasting information from a web site or another source, and then revising it so that it sounds like your original idea.
- Doing an assignment/essay/take home test with a friend and then handing in separate assignments that contain the same ideas, language, phrases, etc.
- Quoting a passage without quotation marks or citations, so that it looks like your own.
- Paraphrasing a passage without citing it, so that it looks like your own.
- Hiring another person to do your work for you, or purchasing a paper through any of the on- or off-line sources.