

University of Florida College of Public Health and Health Professions
PHC7065: Critical Skills in Epidemiology Data Management
Spring 2015
Thursday 6:00 pm to 8:00 pm
Classroom: HPNP G111

Instructor Information

Arbi Ben Abdallah, DES, PhD
Research Assistant Professor
Contact: 314-799-4696 (Cell)
Office: CTRB 4224
Email: aba@wustl.edu
TA: TBA
Office hours: TBA

Course Overview or Purpose

This course focuses on providing basic knowledge and skills needed in data management for research. Included will be: the creation of databases with applications to epidemiological and clinical research; data entry and database management, assessments for data accuracy and consistency (skip patterns, missing data, algorithms, observing data); preparation of project flow logs, codebooks and final databases for statistical analysis; preparation of project data reports; data summarization for presentation (how to read, label and prepare tables and figures), and responsible management of human subject data. The course is designed for advanced students to learn the “code of best practice” for management of research data and to demonstrate proficiency in epidemiological data management before they earn their doctorate.

Course Objectives and/or Goals

The overall goal of this course is to develop competence in data management for epidemiological studies. Specific outcome competencies that are among the goals of this course include, but are not limited to, competency in:

1. The collection and maintenance of epidemiological research data - Students will become familiar with the design and use of the multiple, sophisticated data entry systems and how to “break” them for errors.
2. The development of skills in data manipulation - Students will acquire the skills in data import and export, data transposition, data subset and merging.
3. The preparation of data for statistical analysis - Students will gain the experience in error checking, dealing with missing values, variable format and labeling, variable selection, working with character variables, and date and time variables.
4. Data summarization for professional audience- Students will gain the skills for creating tabular and graphical summaries of epidemiological research data and expected to write brief descriptions of the summaries.
5. The development of codebooks for perpetual use of datasets long after the data are collected.
6. Ethical maintenance to human subject data – Students will learn the important issues regarding data storage and backup, data sharing and transferring, and meta data maintenance.

Prerequisites

PHC 6052 and PHC 6000 or the equivalent.

Course Materials

There is no required textbook for this course.

Recommended Supplemental Textbooks:

Ronald P. Cody and Jeffrey K. Smith. Applied statistics and the SAS programming language, 5th Edition, 2005, ISBN# 01314653325.

Lora Delwiche and Susan Slaughter. The Little SAS Book: A Primer, Fourth Edition, 2008, ISBN# 1599947250. Students may access this book for free through the UF library system.

Course Requirements and Grading

Grades will be based on points accumulated for class attendance & participation, assignments, and final project. Each assignment will be due at the beginning of the following class. A total of 10 points are assigned for each assignment and appropriate amount of points will be deducted for a wrong answer. Total points earned will be assigned as follows:

<u>Component</u>	<u>Percent</u>
Class attendance & participation	30%
Individual & group assignments	50%
Final project/presentation	20%

The grading scale for this course is as follows:

93% - 100%	= A
90% - 92%	= A-
87% - 89%	= B+
83% - 86%	= B
80% - 82%	= B-
77% - 79%	= C+
73% - 76%	= C
70% - 72%	= C-
67% - 69%	= D+
63% - 66%	= D
60% - 62%	= D-
Below 60%	= E

These letter grades translate to the following grade points:

Letter Grade	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E	WF	I	NG
Grade Points	4.0	3.67	3.33	3.0	2.67	2.33	2.0	1.67	1.33	1.0	0.67	0.0	0.0	0.0	0.0

More information about grades and grading policy, please refer to the link:

<http://catalog.ufl.edu/ugrad/current/regulations/info/grades.aspx>

Assignments

Projects will be chosen by each student with the instructors guidance based on the student's research. Projects may include the development of: a codebook for current data, the development of a database for data entry, a cross walk between an assessment and its algorithms, a presentation on a challenging topic in data management, a new statistical method or technology.

In class assignments will be given weekly to reinforce learning.

Topical Outline

Week 1	Understanding the basics of the research process: introducing the principles of research design and data management plan, with emphasis on data collection, processing & storing; use of survey questionnaires, variable names and data types, value range, skip pattern development & handling, double entry and validation.
Week 2	Data import and export: importing and exporting all types of data files such as txt, CVS, Excel, ACCESS, SAS, SPSS and STATA.
Week 3	Data transpose: Reshape data from long to wide format or wide to long format with PROC transpose and array statement using SAS.
Week 4	Data subset and merging: Combine data files, both vertically (appending) and horizontally (merging); validate the data merging; issues when combining data sets with repeats of identifiers.
Week 5	Assessing data quality: Checking errors and missing values; using exploring procedures to check the quality of original and created variables, including frequencies and crosstabs on categorical variables, and descriptive statistics on interval and ratio variables.
Week 6	Getting a "nose" for data; how to observe data and recognize common errors.
Week 7	Data format and cleaning: Variable format and labeling (create formats, or Labels for variable values); variable selection.
Week 8	Working with character variables: Introduction of functions for character variables such as substring; change uppercase or lowercase, removing blank or characters;
Week 9	Working with date and time variables: The formats of date and time; Create date variables; Extract parts from date variable (year, month, weekday or day); interval calculations.
Week 10	Data summarization- CodeBooks
Week 11	Data summarization- Creating Tables
Week 12	Data summarization- Creating Figures and Graphs
Week 13	Ethical maintenance of human health data: data backup and storage; data sharing and transferring; document data management
Week 14	Presentation of final projects
Week 15	Presentation of final projects.

Statement of University's Honesty Policy (cheating and use of copyrighted materials)

Academic Integrity – Students are expected to act in accordance with the University of Florida policy on academic integrity (see Student Conduct Code, the Graduate Student Handbook or these web sites for more details:

<http://www.dso.ufl.edu/sccr/honorcodes/conductcode.php>

<http://www.dso.ufl.edu/studenthandbook/studentrights.php>

<http://gradschool.ufl.edu/students/introduction.html>

Cheating, lying, misrepresentation, or plagiarism in any form is unacceptable and inexcusable behavior.

*We, the members of the University of Florida community,
pledge to hold ourselves and our peers
to the highest standards of honesty and integrity.*

Policy Related to Class Attendance

Class attendance is mandatory. Excused absences follow the criteria of the UF Graduate Catalogue (e.g., illness, serious family emergency, military obligations, religious holidays), and should be communicated to the instructor prior to the missed class day when possible. UF rules require attendance during the first two course sessions (Teachers are required to take roll for the College), and students also must attend all course sessions of student presentations for this class. Missing more than two scheduled sessions will result in a failure. Two weekly sessions are the equivalent of about 15% of the course contact hours. Students should read the assigned readings prior to the class meetings, and be prepared to discuss the material except for the first class session.

Policy Related to Make-up Assignments

Students are responsible for all material presented in class, for meeting the scheduled due dates for class assignments. Personal issues with respect to fulfillment of course requirements will be handled on an individual basis.

Accommodations for Students with Disabilities

If you require classroom accommodation because of a disability, you must first register with the Dean of Students Office (<http://www.dso.ufl.edu/>). The Dean of Students Office will provide documentation to you, which you then give to the instructor when requesting accommodation. The College is committed to providing reasonable accommodations to assist students in their coursework.

Counseling and Student Health

Students may occasionally have personal issues that arise in the course of pursuing higher education or that may interfere with their academic performance. If you find yourself facing problems affecting your coursework, you are encouraged to talk with an instructor and to seek confidential assistance at the UF Counseling & Wellness Center, 352-392-1575. Visit their web site for more information: <http://www.counseling.ufl.edu/>.

The Student Health Care Center at Shands is a satellite clinic of the main Student Health Care Center located on Fletcher Drive on campus. Student Health at Shands offers a variety of clinical services, including primary care, women's health care, immunizations, mental health care, and pharmacy services. The clinic is located on the second floor of the Dental Tower in the Health Science Center. For more information, contact the clinic at 392-0627 or check out the web site at: www.health.ufl.edu/shcc

Crisis intervention is always available 24/7 from:

Alachua County Crisis Center:
(352) 264-6789

<http://www.alachuacounty.us/DEPTS/CSS/CRISISCENTER/Pages/CrisisCenter.aspx>

BUT – Do not wait until you reach a crisis to come in and talk with us. We have helped many students through stressful situations impacting their academic performance. You are not alone so do not be afraid to ask for assistance.