

University of Florida
College of Public Health & Health Professions Syllabus
PHC 7065: Advanced Skills in Epidemiological Data Management (3 credit hours)

Semester: Spring 2026

Delivery Format: On-Campus, HPNP G-110/over Zoom (for students enrolled in the Certificate in
Psychiatric Epidemiology only),

Wednesday: 12:50 pm - 3:50 pm

Course Website: <https://elearning.ufl.edu/>

Instructor Name: Yiyang Liu, PhD MPH

Room Number: HPNP G-110

Office Location: 2004 Mowry Road, Clinical and Translational Research Building (CTRB) #4237

Phone Number: 352-273-5468

Email Address: yliu26@ufl.edu

Office Hours: Tuesday 2:00-3:00 PM (in office CTRB #4237, or over Zoom by request)

Preferred Course Communications: Canvas message, Email

Prerequisites

PHC 6052 and PHC 6000 (or equivalent), or instructor permission.

PURPOSE AND OUTCOME

Course Overview

This course focuses on providing SAS and SQL programming skills for managing and analyzing public health data, including survey and electronic health records. Topics include data import/export, descriptive analysis, recoding, missing data handling, data linkage, relational databases, subqueries, macros, do-loops, and arrays.

Relation to Program Outcomes

This three-credit course is an epidemiology core course for Ph.D. students under the category "Statistics & Data Management." This course will teach students the necessary skills to work with different data formats, perform database management, work with complex relational databases and prepare raw data for different statistical analyses. The course contributes to the Ph.D. program student learning outcomes "Design epidemiologic research studies and analyzes data to answer health-related research questions that are currently relevant to the population."

Course Objectives and/or Goals

The overall goal of this course is to develop SAS programming skills in epidemiological database management. Upon successful completion of the course, students will be able to:

1. Execute data import and export using SAS Wizards and codes; construct and manage SAS libraries.
2. Manipulate numerical, character, and date variables using SAS functions; generate dummy variables.
3. Produce descriptive analyses for continuous and categorical variables.
4. Analyze relational database structures; apply SQL programming skills within the context of

data management; differentiate between various types of SQL joins and unions.

5. Implement SQL programming within SAS and compare its integration with other software tools such as R, Python, and Access.
6. Develop and apply SAS macro variables, functions and programs; design efficient do-loop structures to streamline repetitive coding tasks.
7. Transform datasets between long and wide formats using a variety of methods and evaluate the implications of format choice on downstream analyses.
8. Critically assess the structure and components of Electronic Health Records (EHRs) and biomedical ontologies (e.g., ICD, LOINC, RxNorm), and apply this knowledge to informatic projects in epidemiologic research.

Instructional Methods

We will have in-person class across the 15-week Spring semester. Class begins with a lecture covering basic syntax and key concepts for the week's topic. Students will then work on coding exercises using practice datasets provided on Canvas either at the end of the lecture or between sections. We will review coding solutions together to reinforce understanding and address questions.

What is expected of you?

You are expected to actively engage in the course throughout the semester. You must come to class prepared by completing all out-of-class assignments. This preparation gives you the knowledge or practice needed to engage in higher levels of learning during the live class sessions. If you are not prepared for the face-to-face sessions, you may struggle to keep pace with the activities occurring in the live sessions, and it is unlikely that you will reach the higher learning goals of the course. Similarly, you are expected to actively participate in the live class. Your participation fosters a rich course experience for you and your peers that facilitates overall mastery of the course objectives.

DESCRIPTION OF COURSE CONTENT

Topical Outline/Course Schedule

<i>Date</i>	<i>Agenda</i>	<i>Assignment and Project Due</i>
Jan 14, 2026	Week 1 Course overview	
Jan 21, 2026	Week 2 SAS basic	Assignment 1: posting on Canvas discussion board
Jan 28, 2026	Week 3 Basic SQL SQL in R and Python Special Topic: Health data privacy and classification	
Feb 4, 2026	Week 4 Working with numeric variables	Assignment 2: covers weeks 2 and 3

<i>Date</i>	<i>Agenda</i>	<i>Assignment and Project Due</i>
Feb 11, 2026	Week 5 Working with character and date variables	
Feb 18, 2026	Week 6 SQL and relational database 1 Special Topic: Record linkage methods	Assignment 3: covers weeks 4 and 5
Feb 25, 2026	Week 7 SQL and relational database 2 SQL subquery	
March 4, 2026	Week 8 Hands-on experience working with Electronic Health Record data	Assignment 4: covers weeks 6 and 7 & Project proposal
Mar 11, 2026	Week 9 Midterm Exam	
March 18, 2026	Spring Break, no class	
Mar 25, 2026	Week 10 SAS macro I Macro variable, function, and program Special Topic: Handling missing data 1	
Apr 1, 2026	Week 11 SAS macro II Macro Do Loop Special Topic: Handling missing data 2	
Apr 8, 2026	Week 12 SAS array Transformation between long and wide data format	Assignment 5: covers weeks 10 and 11
Apr 15, 2026	Week 13 Student presentation 1	Final presentation
Apr 22, 2026	Week 14 Student presentation 2	

Course Materials and Technology

Students need to bring their own laptops to all classes and the midterm unless otherwise specified. We will use SAS 9.4 for programming. SAS can be accessed from the PPHP SAS server or UF APPs. There is no required textbook. Here is a textbook that might be helpful to you:

Learning SAS by Example: A Programmer's Guide. Author: Ronald P Cody (2018)

This book is available from UF library ProQuest [E-book Central](#).

For technical support for this class, please contact the UF Help Desk at: Learning-support@ufl.edu, (352) 392-HELP – select option 2, or <https://lss.at.ufl.edu/help.shtml>.

For technical support for this class, please contact the UF Help Desk at:

- helpdesk@ufl.edu
- (352) 392-HELP - select option 2
- <https://helpdesk.ufl.edu/>

For technical support related to PPHP SAS server, please contact PPHP Information Technology (IT) support at:

- support@phpp.ufl.edu
- (352)-273-6200
- <https://it.phpp.ufl.edu/>

Additional Academic Resources

- [Career Connections Center](#): Reitz Union Suite 1300, 352-392-1601. Career assistance and counseling services.
- [Library Support](#): Various ways to receive assistance with respect to using the libraries or finding resources.
- [Teaching Center](#): Broward Hall, 352-392-2010 or to make an appointment 352-392-6420. General study skills and tutoring.
- [Writing Studio](#): 2215 Turlington Hall, 352-846-1138. Help brainstorming, formatting, and writing papers.
- Student Complaints On-Campus: [Visit the Student Honor Code and Student Conduct Code webpage for more information](#).
- On-Line Students Complaints: <https://uf.tfaforms.net/f/ombuds-contact>

ACADEMIC REQUIREMENTS AND GRADING

Assignments

(1) Homework assignments: There are five assignments in total. Assignment 1 is a discussion post about your experience with SAS and other software, and your expectations for the class. Assignments 2 through 5 will be SAS programming tasks that follow the content covered that week. For Assignments 2 through 5, you will need to submit your 1) labeled SAS program in .sas format, and 2) a Word document that includes a copy of your SAS codes, a screenshot of the SAS log, a screenshot of the results (if applicable), and your response (if applicable) for each question. Each assignment has its own point value based on the skills covered, with a total of 40% of the final grade.

Assignment must adhere to the following expectations:

- Your assignment must be turned in no later than 11:59 PM on the due date.

- Late homework assignments without approval will not be accepted (see details in section Policy Related to Make up Exams of Other Work)
- No handwritten assignment. All assignments need to be submitted electronically on Canvas (will be clarified at the beginning of the course).
- For assignments 2-5, you will need to submit a word document that includes the following elements for each question in the assignment:
 - 1) the original question for the assignment
 - 2) copy and paste your SAS codes for each question
 - 3) a screenshot of the SAS log from your submitted code
 - 4) a screenshot of the SAS result output (if specified)
 - 5) your response to the question (if specified)
- You can work with others (e.g., discuss, consult, etc.) on a homework assignment. And, if you work on a homework assignment with other students in the course, you are required to list their names when you turn in the assignment. Directly copying someone else's work and submitting it as your own is not allowed and will be reported as academic misconduct to the Graduate School for investigation.

(2) *Midterm Exam.* Students need to bring their laptops to the classroom to complete the exam. The midterm exam will cover course lectures from week 1 to week 7. The exam will include multiple choices, true/false, or fill in the blank questions. A two-stage assessment will be used for the midterm exam. In the first stage, you will be asked to answer all questions as a closed-book exam. After you submit your exam, you complete the first-stage of the exam and can continue to the second stage. You will not know which questions you got wrong or right after submitting the first stage of the exam. In the second stage, you will be asked to answer exactly the same questions you saw in the first stage. But this time, it will be an open-book exam and you could review all past course materials to help you answer the questions. Your total exam score will be the sum of your score in both stages (the first stage account for 50% of the grade and the second stage will account for the other 50%). Students will have 2 hours to complete the exam.

For example, there are 25 questions in the exam, and each question is worth 4 points. In the first stage of exam, you correctly answer 20 questions. Your score for the first stage of the exam will be 80 points (20*4 points). In the second stage, you reviewed course slides and correctly answer 24 questions. Your score for the second stage of the exam will be 96 points (24*4 points). Your final score for the exam will be $80*50\% + 96*50\% = 88$.

(3) *Final Course Project.* The final project consists of 25% of the final grade and includes 2 parts: a proposal and an oral presentation. This is not a group project and is expected to be completed independently. Students will choose one of the following two options for their final project:

Option 1: Apply SAS skills taught in class to real data

Option 2: Translate SAS to R or Python

Option 1: Student will practice the SAS codes taught in the class with real research data. You will need to apply at least one code from EACH week's class before the midterm (weeks 2 to week 7) and at least one code from ANY courses after the midterm (weeks 10 to 12) to a dataset you are currently working on or a dataset of your interests. You don't need to include any codes from week 1, week 8 (EHR hands-on) and week 9 (midterm exam).

Option 2: This assignment helps you connect the SAS skills you've learned with a programming language you're familiar with (R or Python). You will choose one SAS programming assignment from this class and complete the same coding tasks using R or Python. The goal is to understand how common data management tasks can be completed across different software and syntax.

1. Project proposal requirements:

Submit a one-page proposal and include (50 pts)

- a. Which option do you want pick for your final project (10 pts)

For Option 1:

- b. Descriptions of Data Source you wish to work on: such as what study is the dataset associated with, how it can be accessed, any associated IRB, what are the data formats, and which variables you intend to work with (20 pts)
- c. Which SAS codes you are considering using from each week's lecture before midterm (This could be your tentative plans. You can make changes to it when implementing them) (20 pts)

For Option 2:

- b. Identify the SAS assignment you will translate and whether you will use R or Python. Briefly note which parts will be directly translated, which may need adaptation, and how you will check for equivalent results. (20 pts)
- c. Outline the R or Python code/functions you plan to use for each assignment question. (This could be your tentative plans. You can make changes to it when implementing them) (20 pts)

2. Project presentation requirements:

Prepare a 10 minutes presentation to demonstrate your project followed by a 5 minutes Q&A. (150 pts)

You will be scored based on:

For Option 1:

1. Appropriate application of SAS codes from each required week (60 pts)
2. Code is clearly annotated to show which week's content is being applied, and each step in the data programming process is clearly explained (50 pts)
3. Presentation is well-organized and easy to follow (20 pts)
4. Demonstrates understanding and engagement during the Q&A session (20 pts)

For Option 2:

1. The translation of SAS code into R or Python is clearly explained, showing each step's correspondence to the original SAS assignment and making the logic easy to follow (60 pts)
2. Code is clearly annotated to show correspondence with the original SAS assignment (50 pts)
3. Presentation is well-organized and easy to follow (20 pts)
4. Demonstrates understanding and engagement during the Q&A session (20 pts)

Final project will be evaluated by both the instructor and your peers, with the final grade calculated as 50% instructor score and 50% average peer evaluation score.

Final project in class participation (5%): Each student is expected to actively engage during their classmates' final presentations. Specifically,

1. Verbal participation (3%): you are required to ask one question or provide a thoughtful comment at the end of two different presentations of your choice. Each question or comment will count for 1.5% of your final grade. To receive credit, your contribution must be relevant and demonstrate attention to the presentation content.
2. Peer Evaluation Form (2%): Each student must complete a peer evaluation form for all final project presentations. Completing all forms of each presentation week earns 1% of total grade, for a total of 2% across the two weeks of presentations.

Grading composition

Requirement	Due date	% of final grade (sum to 100%)
Assignments 1-5		Total 40%
Assignment 1	Week 2	4.8%
Assignment 2	Week 4	7.9%
Assignment 3	Week 6	7.9%
Assignment 4	Week 8	9.1%
Assignment 5	Week 12	10.3%
Final Project Proposal	Week 8	5%
Final project presentation	Week 13	20%
Final project participation	Weeks 13/14	5%
Midterm exam	Week 9	30%

- Homework: 40% (5 assignments)
- Mid-term Exam: 30%
- Final project: 30% total; proposal (5%), final presentation (20%), and in class participation (5%)

Point system used

Points Earned	93-100	90-92	87-89	83-86	80-82	77-79	73-76	70-72	67-69	63-66	60-62	<60
Letter Grade	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E
Grade	4.0	3.67	3.33	3.0	2.67	2.33	2.0	1.67	1.33	1.0	0.67	0.0

Please be aware that a C- is not an acceptable grade for graduate students. The GPA for graduate students must be 3.0 based on 5000 level courses and above to graduate. A grade of C counts toward a graduate degree only if based on credits in courses numbered 5000 or higher that have been earned with a B+ or higher

More information on UF grading policy may be found at:

<https://gradcatalog.ufl.edu/graduate/regulations/#Grades>

Exam Policy

The midterm exam will cover course lectures from week 1 to week 7 to demonstrate mastery of the course material. The exam will be completed in-person on Canvas and will include 25 multiple choices, true/false, or fill in the blank questions worth 4 points each. Exams will be given in two phases, the first phase being closed and the second being open-book/note. The exam will be proctored by the instructor and given at during the course period in the same location as the course period.

Policy Related to Make up Exams and Other Work

Make-Up Exam Policy: If you have an expected conflict (e.g., academic conference travel) with scheduled midterm time, you must contact me at least one week before the scheduled exam date to arrange an alternative exam time during an office hour or another approved time. If you miss the exam due to a personal or family emergency, contact me within 24 hours or as soon as possible, to arrange a make-up exam at an agreed upon date.

Late Assignment Submission: The instructor must be informed via Canvas message of any requests for late submission of assignments due to personal/family emergencies within 24 hours of the due date, and preferably before the deadline if circumstances allow. Late assignments due to technical difficulties must also be reported to the instructor within 24 hours of the due date and must be accompanied by the ticket number received from PPHP IT, which serves as the documentation of when the issue arose. Once approved, late assignments should be submitted via Canvas on the original submission page and by the new due date provided by the instructor.

Late submission of assignments without prior permission or for reasons not approved by UF policy will not be accepted and will receive 0 (zero) points.

For university policies on Absence policy, Grading policy, Honesty policy, In-class recording, Gator Evals, please review [UF Academic Policies & Resources](#)

Policy Related to Required 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. Each unexcused absence results in a 1.5% point deduction from the final grade. Late arrivals (arriving more than 20 minutes late) and early departures (leaving more than 20 minutes early) without prior

communication and approval from the instructor will be treated as unexcused absences and will result in a 1.5% point deduction from the final grade. Missing more than three scheduled sessions without excuse will result in a failure. Students are responsible for all material presented in class and meeting the scheduled due dates for class assignments. Please note: There is a separate 5% class participation grade (detailed above), independent of this attendance policy, based on student engagement during final presentations.

Academic Policies & Resources

University academic policies (e.g., Absence policy, Grading policy, Honesty policy, In-class recording, Gator Evals, and etc.) and resources can be found at: [UF Academic Policies & Resources](#)

Policy Related to AI Use in This Course

When authorized by the course director, students may use AI technologies in the completion of coursework as long as they cite all such use by naming the technology and how it was employed. Students assume full responsibility for all content, including errors and omissions. Assistive technology authorized as part of an accommodation for a disability is always permitted.

Course instructors may adjust limitations on AI assistive technology use and must communicate any limitations to students sufficiently in advance of the assignment due date. Failure to cite the use of AI assistive technology, or use of the technology disregarding specific course limitations is considered academic misconduct. The use of AI on assignments, essays/reflection papers, exams, and quizzes when prohibited by course or college instructions is considered cheating and students are violating the UF Regulations 4.040 Student Honor Code and Student Conduct Code.

It is important to note that many generative AI models (e.g. ChatGPT, ChatSonic, Google Bard etc) place any information that they are provided with into the public domain. When using such tools, you must therefore ensure that they are never provided with confidential information. UF AI systems (e.g., Co-Pilot, NaviGator) should never be provided with confidential information. For the avoidance of doubt, the use of such tools is prohibited for generating any confidential communications, including, but not limited to, communications relating to patient records, clients, students and intellectual property. You are also reminded that you should always review the terms and conditions of any third-party software you use (e.g. proof reading tools) to ensure that any data they are provided with is appropriately protected. Always verify information and sources generated by AI tools. AI has been known to generate false information and to cite non-existent sources. Also, because AI-generated text mines people's intellectual property without appropriate credit, this raises ethical concerns.

It is not acceptable to use generative AI for reflective writing, as by its very nature, the process of reflective writing demands that you actively engage in the writing process. Delegating this to a natural language processing algorithm may produce convincing outputs, but does not demonstrate development in your professional practice.

Students are responsible for understanding their dynamic data stewardship responsibilities to minimize personal, college, and university risk.

UF Integrated Risk Management – [CHATGPT Privacy, Factual Accuracy and Usage Guidelines](#)

STUDENT EXPECTATIONS, ROLES, AND OPPORTUNITIES FOR INPUT

Expectations Regarding Course Behavior

Please come to class on time and be prepared to stay until the time scheduled as the end of class. We think your investment in the degree is worth maximizing your in-class experience, and we expect to provide materials that utilize the full, scheduled class times. The use of cell phones is not permitted. Please turn them off or, if you expect urgent calls, set them to “vibrate.” Please do not engage in “side conversations” while the instructor or a presenter is leading the class. If the material is unclear, other students are likely to have a similar question; you are strongly encouraged to ask in-class questions so that all students may benefit from the discussion.

Communication Guidelines

Assistance with course material is available during scheduled office hours or by appointment. Canvas message is the preferred way for communication, and we aim to address all such inquiries within 24 hours of receipt (or on Monday if the email was sent on Friday). Please do not re-send the same question until the appropriate time frame has elapsed (24 hours or end of day Monday for emails sent on Friday). Student success and understanding is of the utmost importance, so each email receives careful consideration.

PHHP Student Resources

PHHP’s UPTurn (Unified Pathways to Support Wellness) program is a *no-cost mental health and wellness program* that is offered year-round to all PHHP students (undergraduate, graduate and professional level) and students (from any college) who are enrolled in PHHP courses. UPTurn advisors support students on their wellness journeys by curating individualized plans (resources and support) to help them manage academic, social, emotional, and health-related stress.

Interested students are paired with an UPTurn advisor, who meets with each student *virtually* (Zoom, Teams, phone) or *in person* (private office/room in HPNP) for a 45-minute consultation, followed by (if desired):

1. Up to 4 follow-up skills-building visits
2. When needed and appropriate, up to 10 psychotherapy sessions after completion of the 4 follow-up skills-building visits

Note: UPTurn is NOT a crisis/emergency resource. Students who are in crisis are strongly encouraged to use UF’s existing crisis support resources, which are listed here:

<https://counseling.ufl.edu/services/crisis/>

Students can learn more about UPTurn and request an appointment here:

<https://phhp.ufl.edu/student-resources/upturn-wellness-program/>

Any questions regarding UPTurn can be directed to upturn@phhp.ufl.edu or (352) 273-6850.

Inclusive Learning Environment

Public health and health professions are based on the belief in human dignity and on respect for the individual. As we share our personal beliefs inside or outside of the classroom, it is always with the understanding that we value and respect diversity of background, experience, and opinion, where every individual feels valued. We believe in, and promote, openness and tolerance of differences in ethnicity and culture, and we respect differing personal, spiritual, religious and political values. We further believe that celebrating such diversity enriches the quality of the educational experiences we provide our students and enhances our own personal and professional relationships. We embrace The University of Florida's Non-Discrimination Policy, which reads, "The University shall actively promote equal opportunity policies and practices conforming to laws against discrimination. The University is committed to non-discrimination with respect to race, creed, color, religion, age, disability, sex, sexual orientation, gender identity and expression, marital status, national origin, political opinions or affiliations, genetic information and veteran status as protected under the Vietnam Era Veterans' Readjustment Assistance Act."