P6110: Statistical Computing with SAS

CLASS SESSIONS
Wednesdays, 5:30pm – 8:20pm
September 7, 2022 – December 14, 2022
Hammer Health Sciences Library – Room LL204

INSTRUCTOR
Debby D'Angelo, MS
dd2542@cumc.columbia.edu

TEACHING ASSISTANTS
Sarah Forrest – sef2183@cumc.columbia.edu
Office Hours: TBD

Xueqing Huang – xh2470@cumc.columbia.edu
Office Hours: TBD

Waveley Qiu – wq2162@cumc.columbia.edu
Office Hours: TBD

Shritama Ray – sr3861@cumc.columbia.edu
Office Hours: TBD

COURSE DESCRIPTION
In this course, you will learn the fundamentals of statistical programming with SAS software. You will begin the course learning the basics of handling datasets in SAS, and then will proceed to learn techniques to manipulate, summarize, and visualize data, and finally perform statistical analyses using appropriate procedures. You will additionally learn how to utilize the SAS macro language to streamline your code, and how to utilize PROC SQL as an alternative to data steps for data manipulation.

COURSE STRUCTURE
This course is formatted as a flipped classroom, which means that lectures are watched on video in advance, outside of class meetings, and "homework" is done during class meetings under the guidance of the teaching team. This format allows you to learn material at your own pace by re-watching lectures as needed, and makes your time spent on exercises more efficient by allowing you to get clarification from the teaching team in real-time as you work.

The beginning of each class meeting will be a review session dedicated to the current week's lecture material that you watched on video prior to class. A weekly quiz will immediately follow each review session (see "Quizzes" under the "Assessment and Grading Policy" section of the syllabus). The remainder of the class session will be spent working on programming assignments. You are encouraged to work with classmates on your assignments; however, you will be responsible for submitting your own work for grading. If you have a question as you work, raise your hand and a TA or the instructor will assist you. Each class meeting will be recorded and posted to Canvas under Echo360 for you to review later.

For questions about course content that come up outside of class meeting times, please send an email addressed to Debby <u>and</u> all TAs, so that whoever is available first can respond promptly. You are also encouraged to use the Discussions board, as your peers can answer and learn from your questions – please include a description, screenshots, or error messages that may help us answer your question. (Your participation in Discussions on Canvas will positively influence your participation grade for the course.) For any personal matters (e.g., missing class or requesting an extension), please email Debby directly.

COURSE LEARNING OBJECTIVES
Students who successfully complete this course will be able to:
- Manipulate data in SAS using data steps and PROC SQL
- Summarize and visualize data using appropriate procedures
- Use SAS macro language to streamline programs
- Choose and perform appropriate statistical analyses for a given dataset and research question, and interpret the results

PREREQUISITES
Students should have taken an introductory biostatistics course (P6103, P6104, Quant Core, or equivalent) and have basic knowledge of statistical inference and regression modeling.

TEXTBOOKS
There are no required textbooks for this course; however, the books below were used in the development of course content and are recommended as supplemental resources:

- Learning SAS by Example: A Programmer's Guide, 2nd Edition by Ron Cody. ISBN-13: 978-1635266597

- SAS Certified Specialist Prep Guide: Base Programming Using SAS 9.4 by SAS Institute. ISBN-13: 978-1642951905

- The Little SAS Book: A Primer, 6th Edition by Lora D. Delwiche and Susan J. Slaughter. ISBN-13: 978-1642952834

SOFTWARE
SAS OnDemand for Academics is available for free for those with a university email. Instructions and links to download here: https://www.sas.com/en_us/software/on-demand-for-academics.html
You may also use the full version of SAS on classroom and CUIMC library computers. NOTE: all examples in lecture demonstrations will be done in SAS 9.4 (the full, paid version of SAS software).

ASSESSMENT AND GRADING POLICY
Final grades will be based on:

In-Class Assignments 30%
Quizzes 20%
Midterm Exam 20%
Final Project 20%
Participation 10%

In-Class Assignments: (30% of final grade) We will cover a total of 15 programming assignments in class. The in-class assignments are designed to give practice implementing the concepts covered in lecture, and to gain experience with troubleshooting issues that may arise while programming. You are highly encouraged to take advantage of the flipped classroom format and seek feedback from peers and the teaching team as you work on assignments; however, **you are expected to submit your own unique work for grading**.

Grading rubrics will be posted simultaneously with the assignment so that you may keep the grading scheme in mind as you work. In addition to a completed grading rubric, you will receive detailed feedback from the TAs to help you improve your skills and clarify concepts that you may have missed in the assignment.

Assignments will be due 3 days following class on **Saturdays at 11:59pm.** Assignments submitted past the deadline without prior approval will have 10% out of the total deducted per day, and if an assignment is not submitted prior to the next class meeting, no credit will be given.

Quizzes: (20% of final grade) You will take weekly quizzes in class to assess your comprehension of the lectures you watched to prepare for the current week. Lecture videos will be unlocked for the following week immediately after class on Wednesdays. Each video contains demonstrations that are designed to help build your skills on various topics. You should follow along with each demo, using the files provided, to prepare for the following week's quiz (and ultimately, in-class assignments).

The quizzes will be administered on Canvas and will be published following a review session at the beginning of class, during which you can ask questions to clarify your understanding of the lecture material. Since the quizzes will happen toward the beginning of class, **please join class on time so that you can benefit from the review and avoid missing the quiz**. No make-up quizzes will be given.

Quizzes will be open-book with approximately 10 questions and will consist of a mix of true/false, multiple choice, fill in the blank, matching, and short answer questions. There will be a **15-minute time limit** for each quiz. **You may not communicate with others while taking the quiz.**

Midterm Exam: (20% of final grade) There will be an open-book midterm exam given during the semester to assess your comprehension of the materials covered in the first half of the semester (Weeks 1-6). The format of the midterm exam will be announced closer to the exam date. **You may not communicate with others while taking the exam.**

Final Project: (20% of final grade) Your final assessment will consist of an extended data analysis using materials covered in this course. Details will be posted toward the end of the semester. This is an individual assessment and **no collaboration is permitted.**

Participation: (10% of final grade) Students who attend class on time, remain for the duration of the class sessions, and ask questions during class can expect to earn full marks for participation. Participating in Canvas discussion boards will affect your participation score favorably. If you expect to be absent or late to class, please let us know in advance.

A+   Reserved for highly exceptional achievement.
A   Excellent. Outstanding achievement.
A-   Excellent work, close to outstanding.
B+   Very good. Solid achievement expected of most graduate students.
B   Good. Acceptable achievement.
B-   Acceptable achievement, but below what is generally expected of graduate students.
C+   Fair achievement, above minimally acceptable level.
C   Fair achievement, but only minimally acceptable.
C-   Very low performance.
F   Failure. Course usually may not be repeated unless it is a required course.

MAILMAN SCHOOL POLICIES AND EXPECTATIONS

Students and faculty have a shared commitment to the School's mission, values and oath.
https://www.mailman.columbia.edu/about/mission-history/public-health-oath

Academic Integrity

Each student in this course is expected to adhere to the Mailman School Honor Code, available online at https://www.mailman.columbia.edu/people/current-students/community-standards/student-honor-code

You are encouraged to utilize the flipped classroom format of this course to teach, and learn from, your peers; however, **all assignments and quizzes must consist only of your own individual work. Only the final project may consist of more than one student's work.**

For assignments, exams, and the final project, you may never submit as your own work:

- Any part of another student's work (including past or current students in any section of this course)
- Any part of a past or current solution file that was posted in class
- Any part of a file obtained from external sources (e.g., internet sites, files created by other programmers, etc.)

During quizzes and exams, you are not permitted to:

- Compare answers with another student, copy answers from another student, or submit answers that you obtained from an external source
- Communicate with anyone other than the teaching team
- Record questions and/or disseminate questions or solutions to others

Additionally, you may not share course materials such as lecture slides, videos, assignments, program files, or any other resources posted to our course's Canvas site with any individual who is not currently enrolled in this course.

Those found to have violated the standards of academic integrity may receive a score of zero on any assessments in question, and may be referred for further disciplinary action.

In order to receive disability-related academic accommodations, students must first be registered with the Office of Disability Services (ODS). Students who have or think they may have a disability are invited to contact ODS for a confidential discussion at 212.854.2388 (V) 212.854.2378 (TTY), or by email at disability@columbia.edu. If you have already registered with ODS, please speak to your instructor to ensure that they have been notified of your recommended accommodations by Meredith Ryer (mr4075@cumc.columbia.edu), Assistant Director of Student Support and Mailman's liaison to the Office of Disability Services.

COURSE SCHEDULE
The following is a tentative course schedule for the Fall 2022 semester. The schedule may change to accommodate the needs of the class.

All course-related documents can be found under the **Files** section of Canvas.
Lecture videos can be accessed under the **Modules** section of Canvas.
Assignments should be submitted through the **Assignments** section of Canvas.

Canvas site: https://courseworks2.columbia.edu/courses/161175

Class Meeting	Topics	Assignments
Week 1 – 9/7/22	Course Introduction, Intro to SAS	Introduce Yourself! Quiz
Week 2 – 9/14/22	Data Manipulation	In-Class Assignment 1 Quiz for Week 2 Lectures
Week 3 – 9/21/22	Summarizing and Visualizing Data	In-Class Assignment 2 Quiz for Week 3 Lectures
Week 4 – 9/28/22	Reporting SAS Output	In-Class Assignment 3 Quiz for Week 4 Lectures
Week 5 – 10/5/22	Efficient Programming	In-Class Assignment 4 Quiz for Week 5 Lectures
Week 6 – 10/12/22	Exporting SAS Output	In-Class Assignment 5 Quiz for Week 6 Lectures
Week 7 – 10/19/22	Midterm Exam	
Week 8 – 10/26/22	Hypothesis Testing for Means	In-Class Assignment 6 Quiz for Week 8 Lectures
Week 9 – 11/2/22	Categorical Data Analysis	In-Class Assignment 7 Quiz for Week 9 Lectures
Week 10 – 11/9/22	Correlation and Linear Regression	In-Class Assignment 8 Quiz for Week 10 Lectures
Week 11 – 11/16/22	Generalized Linear Models	In-Class Assignment 9

		Quiz for Week 11 Lectures
11/23/22	No Class – Thanksgiving Break	
Week 12 – 11/30/22	Survival Analysis	In-Class Assignment 10 Quiz for Week 12 Lectures
Week 13 – 12/7/22	Special Topics 1: Longitudinal Data Analysis, Power and Sample Size Calculations	In-Class Assignment 11 Quiz for Week 13 Lectures
Week 14 – 12/14/22	Special Topics 2: PROC SQL	In-Class Assignment 12 Quiz for Week 14 Lectures
12/19/22	Final Project Due	