DEPARTMENT: BIOS

COURSE NUMBER: 545   **SECTION NUMBER:**

CREDIT HOURS: 2   **SEMESTER:** Spring 2019

COURSE TITLE: Introduction To R Programming

INSTRUCTOR NAME: Ziyi Li

INSTRUCTOR CONTACT INFORMATION

   EMAIL: ziyi.li@emory.edu

COURSE DESCRIPTION

This class is designed to provide an introduction to R and R programming. The objectives of this course are for the student to be able to accomplish the following: Understand the concepts of the R programming language

- Write and debug R functions
- Create useful graphs
- Reshape and aggregate data
- Make adjustments to code to measure and improve performance
- Create Reproducible R Code for distribution and collaboration

This class is for students whose research requires non-trivial amounts of data manipulation and analysis. As with other domains such as Biology and Genetics, Public Health is experiencing a "Data Deluge" which is being addressed in large part by the application of innovative quantitative approaches that require a knowledge of R programming to implement. Therefore, we concentrate on developing programming skills as opposed to data analysis and/or statistical techniques although we do use examples from both these areas to illustrate important concepts. While this is an introductory class, any previous programming experience with other languages will be helpful. This class takes place in the Winter/Spring Semester of 2019. The format will be a mixture of lectures and labs. The schedule below represents the intended content although the instructor may elect to substitute in topics as deemed appropriate.

MPH/MSPH FOUNDATIONAL COMPETENCIES:

- Analyze quantitative and qualitative data using biostatistics, informatics, computer-based programming and software, as appropriate
- Perform effectively on interprofessional teams
- Apply systems thinking tools to a public health issue

CONCENTRATION COMPETENCIES:

- Use statistical software for data management and exploratory data analysis.
- Use statistical software for both data management and data analyses, including coding of custom techniques

EVALUATION

Homework (Four): 80%

Students will be assigned 4 independent homework tasks which will evaluate the student's mastery of the lecture material and topics discussed in class. The homework will present coding tasks of varying complexity and require a correct solution combined with relevant comments and, where required, an R package or Notebook designed to ease reproducibility of the result.

Mandatory Final (In Class, open notes): 20%

The Final exam will be cumulative in nature and questions appearing on the exam will be drawn from material covered and discussed during the semester. The Final exam will also measure the student's ability to write basic software modules on-demand which is an essential skill in R programming.

Grade scale:

Grade	Meaning	Quality Points
A	95-100	4.0
A-	90-94.9	3.7
B+	85-89.9	3.3
B	80-84.9	3.0
B-	75-79.9	2.7
C	70-74.9	2.0
F	<70	0.0

COURSE STRUCTURE

The course will be organized into weekly lectures consisting of a combination of electronic slides and computational demonstrations. Students are expected to ask and answer questions in class.

MPH/MSPH Foundational Competency assessed	Representative Assignment
Analyze quantitative and qualitative data using biostatistics, informatics, computer-based programming and software, as appropriate	Homework assignments and exams will involve the create of functions and packages based on understanding specific data sets
Perform effectively on interprofessional teams	Homework assignments and exams will involve scholarly discussion about approaches to developing solutions without sharing specific code
Apply systems thinking tools to a public health issue	Lectures will emphasize common barriers to developing code. Homework questions will be assigned that require students to develop code and packages that will be assessed and used by other professional in the realm of public health.
BIOS Concentration Competencies assessed	**Representative Assignment**
Use statistical software for data management and exploratory data analysis.	Homework assignments and exams will require programming in R or a similar language.
Explain fundamental concepts of probability and inference used in statistical methodology.	Homework assignments will require interpreting results of a statistical analysis, including interpreting confidence intervals, p-values, Bayesian inference, etc…

COURSE POLICIES

Students are expected to attend lectures and ask questions during class. For computational assignments, students are required, to bring a laptop to class to follow along with code demonstrations. There is no official textbook for this course though the textbooks listed below might be helpful. However, I will discuss the availability of high quality, open source, FREE documentation for R so it is recommended that you wait until after the first session before considering the purchase of any books. Also consider that Emory offers a subscription to the technical series of Safari publications which includes many books on R, Python, and Data Science. All Emory students, faculty, and staff can access this via this link

- R for Beginners, Paradis, E.
- A Beginner's Guide to R, Zuur, Ieno, Meesters
- Applied Epidemiology Using R, Aragon, Thomas
- Software for Data Analysis: Programming with R, Chambers, John

As the instructor of this course I endeavor to provide an inclusive learning environment. However, if you experience barriers to learning in this course, do not hesitate to discuss them with me and the Office for Equity and Inclusion, 404-727-9877.

RSPH POLICIES

<u>Accessibility and Accommodations</u>

Accessibility Services works with students who have disabilities to provide reasonable accommodations. In order to receive consideration for reasonable accommodations, you must contact the Office of Accessibility Services (OAS). It is the responsibility of the student to register with OAS. Please note that accommodations are not retroactive and that disability accommodations are not provided until an accommodation letter has been processed.

Students who registered with OAS and have a letter outlining their academic accommodations are strongly encouraged to coordinate a meeting time with me to discuss a protocol to implement the accommodations as needed throughout the semester. This meeting should occur as early in the semester as possible.

Contact Accessibility Services for more information at (404) 727-9877 or accessibility@emory.edu. Additional information is available at the OAS website at http://equityandinclusion.emory.edu/access/students/index.html

Honor Code

You are bound by Emory University's Student Honor and Conduct Code. RSPH requires that all material submitted by a student fulfilling his or her academic course of study must be the original work of the student. Violations of academic honor include any action by a student indicating dishonesty or a lack of integrity in academic ethics. *Academic dishonesty refers to cheating, plagiarizing, assisting other students without authorization, lying, tampering, or stealing in performing any academic work, and will not be tolerated under any circumstances.*

The RSPH Honor Code states: "Plagiarism is the act of presenting as one's own work the expression, words, or ideas of another person whether published or unpublished (including the work of another student). A writer's work should be regarded as his/her own property."
(http://www.sph.emory.edu/cms/current_students/enrollment_services/honor_code.html)

COURSE CALENDAR AND OUTLINE

Topics and dates are subject to change based on instructors determination of class need

Date	Topic(s)	Supplemental
01/16/18	Motivations, History, Importing Data	
01/23/18	Variables, Objects, Vectors	
01/30/18	Matrices and Factors	
02/06/18	Lists and Data Frames	
02/13/18	Data Frames and Control Statements	
02/20/18	Functions Part 1	
02/27/18	Functions Part 2 and Debugging	
03/07/18	Graphics Part 1	
03/14/18	**Spring Break NO CLASS**	
03/21/18	Graphics Part 2	
03/28/18	Package Development Part 1	
04/04/18	Package Development Part 2	
04/11/18	Aggregation Using dplyr	
04/18/18	Statistical Analysis	
05/02/18	Git	
05/09/18	Final (In class)	