Syllabus Module Module 215: Minor B "Introduction to *R*: computing, graphics and statistics"

Module : 215	Module 215: Minor B "Introduction to *R*: computing, graphics and statistics"
UE coordinator	Charles DiMaggio, PhD, MPH Professor of Surgical Sciences and Population Health, Director of Injury Research New York University School of Medicine 462 First Avenue, NBV 15 New York, NY 10016-9196 Charles.DiMaggio@nyumc.org
Dates	Monday 24th – Friday 28th October 2016
Credits/ECTS	3ECTS
Duration	30 hours (5 days of 6h)
UE description	The course will teach students to apply basic analytical epidemiologic methods and data manipulation using the R statistical computing. The emphasis will be on learning by doing. Practice material and data (available on the instructor's website) will be grounded on actual research questions, often based on the instructor's own work, and are intended to illustrate the kinds of issues that often arise when practicing epidemiology. Students will be encouraged to apply the material to their own research interests and epidemiological analyses. Each session is a combination of didactic lecture and hands-on practice. Students will conduct epidemiological analyses on actual data sets, and learn the importance of data preparation and cleaning, descriptive analyses, as well as how to conduct basic statistical analyses including regression analyses.
Prerequisites	Students are assumed to be familiar with elementary statistical methodology (such as regression models, analysis of variance, hypothesis testing, etc) and computer science
Course learning objectives	The course will familiarize students with the programming skills necessary to: • Use the R statistical computing environment to enter, read, clean, organize, index and manipulate epidemiologic data in R. • Analyze rate, ratio and risk data in R using using matrices, arrays and data frames. • Apply R functions and packages for linear models to analyze epidemiologic data. • Apply base graphics and ggplot2 to explore, plot and present analyses. • Understand how R can be extended with packages like rgdal, sp, and maptools to conduct basic spatial analyses and data mapping.

UE Structure
(details of sessions title/spaeker/date/duration)
	Class meets from 9:00 am – 12:00 pm and 2:00 pm – 5:00 pm Monday – Friday for a week (30 hours total). We will typically begin each session with a structured lecture followed by hands on applications of course material. All sessions are delivered by Dr Charles Di Maggio, details of each session are provided below.
	
	• Getting started in R, *Day 1, AM*
	    Installing, Calculating, Entering Data
	• Components of the R Computing and Statistical Analysis Platform, *Day 1, PM*
	    Examples of Analyses, Introduction to R Objects
	• Objects and Data, *Day 2, AM*
	    Vectors, Matrices, Lists, Dataframes
	    Introduction to Indexing
	• Functions and Indexing, *Day 2, PM*
	    Marginal and Stratified Analyses, Summary Statistics
	    Using Indexing to Manipulate Data
	• Statistics in R, *Day 3, AM*
	    Statistical Functions, t-tests, F test, boxplots, non-parametric Tests,
	    Introduction to Regression Functions

	• Regression in R, *Day 3, PM*
	   Linear, Logistic and Poisson Regression
	• Epidemiologic Functions in R*, Day 4, AM*
	   Relative Rates and Risks, Survival Analysis
	• Introduction to SAS and SAS procedures,*Day 4, PM*
	   Graphics in R
	• Special Topics, Course Projects*, Day 5, AM*
	• Course Projects, Discussion, Wrap Up*, Day 5, PM*

Course requirement	Students have to come to class with a computer with the software R installed, and a flash drive with at least 500 MB of memory.
Grading and assessment	In-class Exercises: 20% A set of introductory exercises and practical tutorial material will be made available for download from the instructor's website. Students may complete the material at their own pace and submit for grading 1 week after the end of class. Final Project: 80% Students are required to make a short presentation and submit a final paper of an analytic vignette. Students may work in groups of up to 3. The assignment is intended to allow the student the opportunity to demonstrate his or her ability to conduct preliminary analyses of an epidemiologic problem using R, and to share that information with colleagues. The vignette assignment will be a combination of a brief overview of an epidemiologic question, ideally related to the student's research interest, followed by a narrative discourse of an analytic approach to answer that question uisng R. Actual or simulated should allow that code to be run. Students will present their vignettes as part of oral presentations during the last day of the class, and submit the written assignment 1 week after the end of class. The final paper is limited to 3-pages, double-spaced with 1-inch margins. It should be The 5-page final assignment consists of: • Cover Page: Title, student's same, Affiliations (if any) and abstract or overview of the vignette including methods and results. • Introduction: What is known about the epidemiological topic, what is not known, why it is important. • Approach: Analytic design or method to be presented, including variables and their roles in the analysis (explanatory, confounding, interaction, etc...). The R package to be explored including a brief listing or description of the functions to be presented. • Implementation: Narrative walkthrough of the analysis including full code and pilot or synthesized data to allow reproduction. • Discussion: Including limitations, conclusions, possible next steps. • References: Limited to 15 citations (not counted toward page limit) Late submission of the final assignment will result in a penalty of 10 points for each day the assignment is late.
Location	George Sand EHESP Campus in Paris
Readings	Required material is posted at: http://www.injuryepi.org/styled-4/styled-6/ The following texts are recommended: • An Introduction to R. Free at http://cran.r-project.org/doc/manuals/R-intro.pdf (also comes with default installation) • Aragon TJ, et al. Applied Epidemiology Using R. Free at http://www.medepi.com/