ERTH 615

LITERATE PROGRAMMING WITH R, RSTUDIO, AND R MARKDOWN

PURPOSE

This **3-credit** course was created to meet the needs of graduate students and upper level undergrads who want to improve their programming skills. **R** is the freeware, open-source version of the S computer language originally created at Bell Labs expressly for interactive data analysis. It has a colossal user base of data analysts, and over 10,000 special packages of functions for data analysis. **RStudio** is a free interactive development environment (IDE). **Literate programming** is the paradigm in which the code and narrative exist in the same document. We use **RMarkdown**, which allows one to embed not only R, but also Python, Stan and many other computer languages. (RMarkdown is what Jupyter notebooks aspires to be when it grows up.)

This course is partly lecture, but mostly hands-on coding. Weekly coding assignments cover basic data analysis, some basic calculus, some linear algebra (vectors, matrices) and some standard "big data" techniques such as principal components and spectrograms.

Instructor: Neil Frazer, neil@hawaii.edu, POST Building Room 819-C.

Times and place: This course will be **in-person with social distancing** in a large, laboratory room with tables and comfortable chairs: Tuesday & Thursday 12:00-1:15 in POST Building, Room 703

Syllabus

Numbered roughly by week, although the order of topics may change.

1. Introduction to R, RStudio and R Markdown. Simple computations in R. YAML headers, cascading style sheets, chunks and the setup chunk. Knitting to html, MSWord and PDF.

2. Introduction to graphics in R. LaTeX for math in R Markdown. Expressions and plot annotation in R.

3. Classes and data structures in R. Vector and matrix algebra review.

4. Indexing and subsetting. Big data skill: principal components.

5. Skill with RStudio: navigating directories and environments, inspecting and editing data structures, choosing and installing packages, CRAN and TaskView. Skill with R Markdown: stopping the knit where you want, figure sizing, choosing the dev, saving separate figures.

6. Writing and calling functions. Graceful error handling. Big data skill: spectrograms.

7. For- and while loops. Functional programming. Advanced graphics: legend wrangling.

8. Control structures: if, ifelse, if-else, and switch. Review of probability density functions.

9. Forward modeling and the estimation of uncertainty. Why Monte Carlo techniques such as the bootstrap are easier and less prone to error than calculus. Histograms, densities and statistics. Why quantiles are usually more robust and informative than the mean and standard deviation.

10. The dpqr of distributions: PDFs, CDFs, Complementary CDFs (survival functions), inverse CDFs (quantile functions) and samplers. Types of measure: ratio scale, interval scale and proportion. When to transform.

11. Frequentist thinking and likelihood. Regression. Confidence intervals and prediction intervals. Robust methods for noisy data. Bootstrapping.

12. Bayesian thinking: prior, posterior and sampling distributions, likelihood, prior predictive, posterior predictive, evidence. Conjugate priors. Uninformative priors.

13. Introduction to Markov chain Monte Carlo (MCMC).

14. Review.

Textbooks: No text is required. **Prerequisites**: Intructor permission.

Assessment and Grading:

Literate programming, no less than traditional programming, requires regular practice, so this course utilizes weekly coding assignments that must be handed in by the due date except for medical reasons. Most students will be able to complete the assignments during the laboratory period. The final grade will be a weighted average of attendance (10%), scores on weekly assignments (80%), and scores on occasional quizzes (10%). There is no final exam.

Class Format:

This is a combined lecture-laboratory. Students are encouraged to actively ask questions in class, to work on weekly assignments in class, and to assist each other in learning, although copying of work is not permitted.

ERTH Learning Objectives:
The Earth Sciences department has defined 4 and 5 student learning objectives (SLOs) for the MS and Ph.D. degree programs, respectively. For the MS they are (1) Technical knowledge, (2) Scientific method, (3) Communicate Earth Sciences knowledge, (4) Employability/Contributions Post-Graduation. For the Ph.D., an additional SOL is required, i.e., (5) Expertise in a sub-discipline. This course directly incorporates content relevant to the following SLOs:

- SLO 1: Through practice with R, RStudio and R Markdown, students will learn the techniques and strategies common to all scientific programming languages.
- SLO 2: Students will acquire a basic understanding of data analysis.
- SLO 3: Students will learn to produce accurate, readable reports that incorporate code as well as narrative.
- SLO 4: Students will acquire expertise in a language widely used in industry as well as science.
- SLO 5: (PhD only) Students will learn non-parametric data analysis techniques such as the bootstrap.

Disability Access:

If you have a disability and related access needs the Department will make every effort to assist and support you. For confidential services students are encouraged to contact the Office for Students with Disabilities (known as "Kokua") located on the ground floor (Room 013) of the Queen Lili'uokalani Center for Student Services:

KOKUA Program; 2600 Campus Road; Honolulu, Hawaii 96822
Voice: 956-7511; Email: kokua@hawaii.edu ; URL: http://www.hawaii.edu/kokua

Discrimination:

The University of Hawai'i is committed to providing a learning, working and living environment that promotes personal integrity, civility, and mutual respect and is free of all forms of sex discrimination and gender-based violence, including sexual assault, sexual harassment, gender-based harassment, domestic violence, dating violence, and stalking. If you or someone you know is experiencing any of these, the University has staff and resources on your campus to support and assist you. Staff can also direct you to resources that are in the community. Here are some of your options:

As members of the University faculty, instructors are required to immediately report any incident of potential sex discrimination or gender-based violence to the campus Title IX Coordinator. Although the Title IX Coordinator and your instructors cannot guarantee confidentiality, you will still have options about how your case will be handled. Our goal is to make sure you are aware of the range of options available to you and have access to the resources and support you need.

If you wish to remain ANONYMOUS, speak with someone CONFIDENTIALLY, or would like to receive information and support in a CONFIDENTIAL setting, use the **confidential resources available here**: http://www.manoa.hawaii.edu/titleix/resources.html - confidential

If you wish to directly REPORT an incident of sex discrimination or gender-based violence including sexual assault, sexual harassment, gender-based harassment, domestic violence, dating violence or stalking as well as receive information and support, contact: Dee Uwono Title IX Coordinator (808) 956-2299 t9uhm@hawaii.edu.