STA 3100 (17740 and 27387)

Spring 2025

Programming with Data (in `R`)

Class 17740 MWF: 10:40-11:30 in FLO 100

Class 27387 MWF: 09:35-10:25 in WM 100

Instructor: Demetris Athienitis

Office: Griffin Floyd 116B
E-mail: athienit@ufl.edu

Teaching Assistants:

Zhaoyan Song
E-mail: zhaoyan.song@ufl.edu

Ziqian Yang
E-mail: zi.yang@ufl.edu

Course Website: e-Learning

Course Delivery: Teaching will be in-person and in-class. Material will be available in e-Learning.

Course Description: An introduction to statistical computing and programming with data. Topics include basic programming in R; data types and data structures in R; importing and cleaning data; specifying statistical models in R; statistical graphics; statistical simulation using pseudo-random numbers; reproducible research and the documentation of statistical analyses.

Prerequisite(s): STA 3032 (B-) or STA 2023 (B) or AP Statistics (4)
Credit Hours: 3

Communication:

- Discussion forum (link to Microsoft Teams available in course website).

- Office hours (posted in e-Learning).

- E-mail for questions regarding course policies. (Ensure that **STA 3100** is in the subject line. Failure to do so may result in a non-response.)

Course Goals and Objectives

1. Import data into R and prepare the data for analysis

2. Write functions in R making effective use of data structures and control structures

3. Formulate statistical models in the R language

4. Perform, document, and interpret common statistical analyses

5. Carry out statistical/probabalistic simulations

6. Determine graphics appropriate to a statistical analysis and produce them using R

7. Document and report the results of data analyses and simulations in a reproducible way

Materials and Supply Fees

All student materials used in this course are free. The course will not use a textbook, instead following notes prepared by the instructor and other faculty. The course will also make heavy use of free software and online materials. The volume of how-to guides for all sorts of projects in R is quite extensive, and the course will sample many of these as is seen fit. Some of the most popular, streamlined, and thorough sources for learning R programming are listed below, but many other very good sources are available.

- r4ds: R for Data Science: Visualize, Model, Transform, Tidy, and Import Data

- r4ds2e: R for Data Science (2e): Import, Tidy, Transform, Visualize, and Model Data

- rp4ds: R Programming for Data Science

- hopr: Hands-On Programming with R : Write Your Own Functions and Simulations

- ggplot2: Elegant Graphics for Data Analysis (3e)

- advr: Advanced R (2nd Ed)

Computational Requirements: Students will need to have frequent and reliable access to a computer capable of running R code. All computers supplied by UFIT Academic Technology classrooms and laboratories, including those at the university libraries, have R and RStudio installed and perform well enough to run any program of interest to this course without any trouble. However, because of data security policies, it's somewhat tedious (though possible) to efficiently manage an R installation on university computers. For these reasons, students with personal computers are encouraged to use them for programming tasks. Students who wish to use personal computers will likely have no trouble with the computational requirements necessary for R and RStudio. You will need the following to install R and RStudio:

- For Windows users, Windows 10 or later.

- For macOS users, macOS 10.13 (High Sierra) or later.

- Other hardware requirements (e.g. RAM, processor speed) will be satisfied if you meet the operating system requirements.

On Windows, R requires about 179 MB of space, and RStudio requires about 861 MB. Other operating systems are likely to have similar storage requirements. Additional storage will be needed for downloaded packages, datasets, and output that students produce. Students should have at least 5 GB of storage available once you have downloaded R and RStudio. Please contact the instructor if you have any concerns about the technology requirements of the course.

Course Policies

The instructor reserves the right to update some parts of this syllabus as necessary.
Students will promptly be notified of any changes.

Demeanor

All members of the class are expected to follow rules of common courtesy in all classroom discussions, email messages, threaded discussion and chats. Please refer to expected class netiquette online and during class.

Assignments

- Students are expected to **show and explain their work**.

- Students are allowed to consult any online resources that are deemed relevant, and may consult with other students. However, the work submitted by each student must be **uniquely** theirs. **Offering** and **accepting** solutions - even from textbook - is an act of **plagiarism**, which is a serious offense and **all involved parties will be penalized according to the UF Student Honor and Conduct Code**. When in doubt, direct your questions to the instructor.

- All **deadlines (excluding exams) are at 23:59** of the due/end date and posted on course page. These are *hard* deadlines meaning that any open or ongoing assignments will automatically be submitted at the deadline. Assignments are automatically submitted at deadline even if in progress.

- Electronically submitted work must contain two files:

  - the compiled pdf file,
  - the source file/script file, e.g. R Markdown .rmd that can be recompiled by the instructor/TA.

  In Canvas, all uploaded files automatically get a grade of 0, until the teaching assistant grades them.

- Feedback will provided within two to five business days from the assignment deadline.

Exams

The (in-class) exams may comprise of multiple choice questions and a few open-ended questions.

Important dates:

(Subject to change)

Exam #1 .... February 19
Exam #2 .... April 21

Grading

Grade distribution:

2 Exams	10% each	
10 Assignments	8% each	
Total	100%	
Extra Credit	0-1%	(class and discussion forum participation)

Letter grade assignment:

There will be *no rounding up* of scores.

		A	91 to 100	A-	88 to < 91
B+	84 to < 88	B	80 to < 84	B-	77 to < 80
C+	74 to < 77	C	70 to < 74	C-	67 to < 70
D+	64 to < 67	D	60 to < 64	D-	55 to < 60
E	< 55				

- Final grades are not shown on e-Learning as they do not account for the conditional weighing of exams.

- A minimum grade of C is required for any programs within the Department of Statistics, i.e. majors/minors.

- To view the result of the letter grades to your GPA please visit the UF Grade and Grading Policies.

Make-up

Requirements for class attendance and make-up exams, assignments, and other work in this course as well as policies regarding absences, religious holidays, illness and student athletes are consistent with UF Attendance Policies and Examination Policies.

Addressing Issues

Technical difficulties

Please contact the UF Help desk via e-Learning "Help" tab. Any requests for make-ups due to technical issues must be accompanied with appropriate documentation/proof including screenshots and communication with the help desk. You MUST contact your instructor within 24 hours of the technical difficulty if you wish to request a make-up.

Grievances/Commendations

Should you have any grievances or commendations with your experience in this course you can always address them to the instructor (anonymously) or to the Department of Statistics. For issues that are not satisfactorily resolved at the department level or which seem to be broader than one department, students are referred to the Office of the Ombuds.

UF and CLAS Policies

Dropping, Withdrawing and Incomplete

Dropping and Withdraw

For late course drops and course withdrawals check the catalog.

Incomplete

An incomplete grade may be assigned at the discretion of the instructor as an interim grade for a course in which the student has completed a major portion of the course with a passing grade, been unable to complete course requirements before the end of the term because of extenuating circumstances, and obtained agreement from the instructor and arranged for resolution of the incomplete grade in the next term. Instructors are not required to assign incomplete grades. For complete details please visit CLAS incomplete grade policy and contract.

Accommodating Students with Disabilities

Students requesting accommodation for disabilities must first register with the Disability Resource Center (DRC). The DRC will provide documentation to the students who must then provide this documentation to the instructor when requesting information. You must submit this documentation prior to submitting any assignments for which you are requesting accommodation.

U Matter, We Care

U Matter, We Care, through the Dean of Student's Office, offers care related resources and programs focused on health, safety, and holistic well-being.

Academic Misconduct

Students are held accountable to the UF Student Honor and Conduct Code.

Evaluations

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at `https://gatorevals.aa.ufl.edu/students/`. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via `https://ufl.bluera.com/ufl/`. Summaries of course evaluation results are available to students at `https://gatorevals.aa.ufl.edu/public-results/`.

Tentative Course Outline

Week	Content
1	Syllabus, RStudio and R basics
2	Objects, data types, data structures, functions, packages
3	Reading datasets, basic dataset properties
4	Basic graphics and plots, basic dataset manipulation.
5	Dataset manipulation, control structures
6	Introduction to RMarkdown
7	Introduction to LaTeX
Exam 1	
8	RMarkdown continued, introduction to the `tidyverse`
9	SPRING BREAK
10-11	Tidyverse data operations.
12	tidyverse graphics, introduction to statistical modeling
13-14	Statistical modeling
15	Miscellanea: data pulling operations, regular expressions
Exam 2	