



EFB 654:

Intro to R and Reproducible Research

Course Overview:

This course serves as an introduction to the use of R, a statistical programming language and environment, for scientific research. R is an open-source tool especially beneficial for processing and analyzing large datasets, an issue present in many scientific fields.

Instructor:

Nicole Barbour, Ph.D.

barbour@esf.edu

Office: Illick 244

This course is designed for beginner to intermediate users of R. In addition to teaching students R basics and essentials, such as syntax, script construction, and use of R packages, this course will also introduce students to using: 1) RStudio for optimal visual coding, 2) RMarkdown to embed code and text into a text-based (e.g., html) document, and 3) GitHub for version control and cloud-based storage of code.

Students will have the opportunity to design and conduct their own final project for the course, an excellent opportunity for those with thesis or capstone data in need of analysis.

Learning Objectives:

- Be able to bring data (such as Excel files, shapefiles, etc) *into* R and be able to export it *out* of R
- Be able to write clean, efficient, and reproducible R code scripts
- Understand how to deal with both basic and advanced R data classes, such as timestamps and spatial data
- Be able to summarize and plot data to illustrate descriptive statistics, visual patterns/trends, or statistical model results
- Be able to construct or “knit” R markdown files that embed both code and text
- Be able to write basic functions, conditional statements (e.g., “if...else...”), and for-loops
- Be able to use base R functions to conduct and interpret basic statistical tests, such as a T-test, Chi Square, ANOVA, or linear regression
- Be able to use Git and Github for version control of R code and Rmarkdown documents
- Be able to apply learned course content to design and carry out a full independent research project that satisfies all of the previously stated Learning Objectives

Class Schedule

M/W 10:35-11:30 am (Baker 309)

Office Hours: W 11:30-2:30 (Illick 244, or by appointment)

Prerequisites

There are no prerequisites. Basic statistics and competency with Excel would be helpful but not required.

Course Content

All course content will be available on Blackboard (<https://blackboard.syr.edu/>), including assignments, example code, and announcements.

Reading Materials and Resources

We will be using the open-source, online books, “*R for Data Science*” (<https://r4ds.had.co.nz/>) and “*R for Fledglings*” (<https://www.uvm.edu/~tdonovan/RforFledglings/>) for supplementary reading and exercises each week.

Some other helpful resources we may use this semester:

- R news and tutorials: <https://www.r-bloggers.com/>
- Spatial Data in R (sf package): <https://r-spatial.github.io/sf/articles/sf1.html>
- A Handbook of Statistical Analysis in R:
<https://vcarrion.people.uic.edu/handbookrstarters.pdf>
- Advanced Interactive Mapping (mapview package):
https://r-spatial.github.io/mapview/articles/mapview_02-advanced.html
- Stack Overflow for R forums and help: <https://stackoverflow.com/>
- Happy Git and GitHub for the useR online book: <https://happygitwithr.com/>
- SwirlR tutorial: <https://swirlstats.com/>

Required Software

You will need to install R (<https://cran.r-project.org/>), RStudio (<https://posit.co/download/rstudio-desktop/>), and Git (<https://github.com/git-guides/install-git>). If you have a Mac, please follow these instructions to download R and RStudio: <https://teacherscollege.screenstepslive.com/a/1135059-install-r-and-r-studio-for-mac>.

Please come to the first class with R and RStudio installed on your computer (or at least, attempted).

Assessments and Grade Breakdown

Student performance and understanding in the course will be assessed using in-class exercises, regular homework assignments, and a final project, consisting of a presentation and written report. In-class participation will be assessed by students submitting in-class exercise answers on Blackboard. Homework assignments will be graded based on completion. Accuracy of homework answers will be self-assessed according to homework solutions provided by the instructor immediately following homework due dates. Presentations and reports will be assessed according to a rubric, provided well within advance of due dates.

Category	Number	Points Each	Total Points	Weight
Participation in Class Exercises	16	5	80	22%
Homework Assignments	9	15	135	37%
Final Project Presentation	1	50	50	14%
Final Project Report	1	100	100	27%
		Total	365	100%

Attendance and Campus Policies

Except for approved/excused absences (illness, death in the family, religious holiday, official University business), attendance to class is mandatory. If you need to miss class (either excused or unexcused), please contact me ASAP to determine how you can make up for any missed content.

The student is responsible for contacting the instructor in a timely manner to discuss any late assignments/missed work/missed lecture content. There is no SUNY ESF college policy related to missed work. This decision belongs to the instructor. If granted extensions/make-ups/incompletes, it is the responsibility of the student to complete the work in the agreed upon timeframe.

Clery Act/Title IX Reporting

I expect all students to respect each other with regards to self-identity, including gender, race, class, sexuality, religion, and disability.

SUNY ESF is committed to enhancing the safety and security of the campus for all its members. In support of this, faculty may be required to report their knowledge of certain crimes or harassment. Reportable incidents include harassment on the basis of sex or gender prohibited by Title IX and crimes covered by the Clery Act. For more information about Title IX protections, go to the Title IX website, or contact the Title IX Coordinator, Rebecca Hoda-Kearse, 220 Bray Hall, titleix@esf.edu, 315-565-3012.

For more information about the Clery Act and campus reporting, go to the University Police Annual Report.

Learning and Physical Disabilities

SUNY-ESF works with the Center for Disability Services (CDR) at Syracuse University, who is responsible for coordinating disability-related accommodations. CDR is responsible for coordinating disability-related academic accommodations and will work with the student to develop an access plan. Since academic accommodations may require early planning and generally are not provided retroactively, please contact CDR as soon as possible to begin this process. To discuss disability-accommodations or register with ODS, please visit their website at <https://disabilityresources.syr.edu>. Please call (315) 443-4498 or email CDRaccommodate@syr.edu for more detailed information.

Academic Dishonesty

Academic dishonesty is a breach of trust between a student, one's fellow students, or the instructor(s). Examples of academic dishonesty include but are not limited to plagiarism and cheating, and other forms of academic misconduct. By registering for courses at ESF you acknowledge your awareness of the ESF Code of Student Conduct. More information regarding Academic Integrity, including the process for resolving alleged violations, can be found in the Student Handbook.

Inclusive Excellence Statement

As an institution, ESF embraces inclusive excellence and the strengths of a diverse and inclusive community. During classroom discussions, all of us may be challenged by ideas different from our lived experiences and cultures. Understanding individual differences and broader social differences will deepen our understanding of each other and the world around us. In this course, all people (including but not limited to, people of all races, ethnicities, sexual orientation, gender, gender identity and expression, students undergoing transition, religions, ages, abilities, socioeconomic backgrounds, veteran status, regions and nationalities, intellectual perspectives and political persuasion) are strongly encouraged to respectfully share their unique perspectives and experiences. This statement is intended to help cultivate a respectful environment, and it should not be used in a way that limits expression or restricts academic freedom at ESF.

Religious Holiday Observance

All students have a right under NYS law and ESF college policy to observe the religious holidays of their choice, according to their individual faith. If students wish to observe a religious holiday, they should provide written notification to the instructor of their intent to observe a particular religious holiday within the first two weeks of the semester, and prior to missing any required course meetings or activities. Reasonable requests for absence from course meetings or activities will be accommodated whenever possible, though students may be responsible for independently making up missed materials or activities on their own time, and in a timely fashion.

Weekly Course Schedule

Content	Date	Readings/Assignments
Class Introduction - <i>What is R and RStudio?</i> Student and Professor Introductions - <i>Why are you taking this class?</i>	Jan 17	R for Fledglings: Chap 1-2 Class Exercise 1 Homework Assignment 1 (Fri, 11:59 pm)
File Organization - <i>Installing packages, File paths and file management</i>	Jan 22	R for Fledglings: Chap 5, 12 Class Exercise 2
R Basics - <i>Data types, Object types, Vectors, Matrices, Arrays, Lists, Data Frames</i>	Jan 24	R for Fledglings: Chap 4 Class Exercise 3 Homework Assignment 2 (Fri, 11:59 pm)
R Markdown - <i>Chunk names, Chunk options, Global options</i>	Jan 29	R for Data Science: Chap 27 Class Exercise 4
Data Wrangling - <i>Creating scripts, Importing/exporting data, Dates and times, Merging/binding/subsetting/joining data, Using NA values, Using dplyr and pipes</i>	Jan 31	R for Fledglings: Chap 6-8 Class Exercise 5 Homework Assignment 3 (Fri, 11:59 pm)
Summarizing What We've Learned - <i>Create an R Markdown file using imported data and dplyr</i>	Feb 5	Class Exercise 6
Functions in R - <i>Base R functions, Writing your own functions, Lists and lapply, Conditional statements, For-loops</i>	Feb 7	R for Fledglings: Chap 3, 11 R for Data Science: Chap 19 Class Exercise 7 Homework Assignment 4

		(Fri, 11:59 pm)
Summarizing What We've Learned - <i>Create an RMarkdown file using your own function, a for-loop, an if-else statement, and lapply</i>	Feb 12	Class Exercise 8
Final Projects Planning Day - <i>Bring or find your datasets, propose research question and methods, plan out workflow</i>	Feb 14	Homework Assignment 5 (Fri, 11:59 pm)
Data Visualization - <i>Base R plots, ggplot2 plots, plot types</i>	Feb 19	R for Data Science: Chap 3 Class Exercise 9
Spatial Data in R, I - <i>Simple features and the sf package, Geometry types, Coordinate systems</i>	Feb 21	Spatial Data in R (online handout) Class Exercise 10 Homework Assignment 6 (Fri, 11:59 pm)
Spatial Data in R, II - <i>Geometrical operations, Plotting spatial data, Interactive mapping with mapview</i>	Feb 26	Spatial Data in R (online handout) Class Exercise 11
Summarizing What We've Learned - <i>Create an RMarkdown file performing geometrical operations, ggplots, and interactive maps of spatial data</i>	Feb 28	Class Exercise 12 Homework Assignment 7 (Fri, 11:59 pm)
Review and Final Project Work Day - <i>Homework review, Discussion period, Final project planning and help</i>	Mar 4	
Git and Github - <i>Installing and enabling Git, Creating a Git project, Committing/pushing/pulling, Git terminal commands</i>	Mar 6	Happy Git Online Book and Tutorial Class Exercise 13 Homework Assignment 8 (Fri, 11:59 pm)
No Class - Spring Break	Mar 11	
No Class - Spring Break	Mar 13	
Summarizing What We've Learned - <i>Create a Git Repo for your final project, Push an</i>	Mar 18	Class Exercise 14

<i>RMarkdown file with text, code, plots, maps</i>		
<i>Statistics in R - Descriptive stats, Basic statistical tests, Building models, Model diagnostics and predictions, Plots and tables of results</i>	Mar 20	R for Fledglings: Chap 9 R for Data Science: Chap 23-24 A Handbook of Statistical Analyses in R online handout Class Exercise 15 Homework Assignment 9 (Fri, 11:59 pm)
<i>Summarizing What We've Learned - Push an RMarkdown with descriptive stats, statistical model results, plot and table results</i>	Mar 25	Class Exercise 16
<i>Review and Final Project Work Day - Homework review, Discussion period, Final project planning and help</i>	Mar 27	
<i>Review and Final Project Work Day - Homework review, Discussion period, Final project planning and help</i>	Apr 1	
<i>Review and Final Project Work Day - Homework review, Discussion period, Final project planning and help</i>	Apr 3	
Final Project Presentations	Apr 8	
Final Project Presentations	Apr 10	
Final Project Presentations	Apr 15	
Final Project Presentations	Apr 17	
<i>Final Project Help Day - Final project help - last chance before reports are due!</i>	Apr 22	
<i>Final Project DUE -----</i>	Apr 24	