Instructor: Yanan Wu TA: Khadija Nisar

Spring 2025



YANAN WU – VISITING ASSISTANT PROFESSOR

Education & Experience



• Python Programming



- 1. Manipulating Spatial Data
- 2. Web Mapping
- 3. Processing Raster
- 4. Data Analysis
- 5. Creating Custom Tool
- 6. Data Visualization
- 7.

...

• Spatial Database



- 1. Geodatabase
- 2. SQL
- **3.** Proximity Analysis
- 4. Geometry processing
- 5. Raster processing
- 6. PostSQL with python
- 7. ...

• Intermediate Statistics



- 1. Bivariate regression
- 2. Logistics regression
- 3. PCA
- 4. GWR
- 5. Spatial Autocorrelation
- 6. ...

INSTRUCTOR OFFICE HOUR

- Instructor: Yanan Wu
- Email: yanawu@clarku.edu
- Office Hours: Tuesday & Thursday 1:30 pm 2:30 pm or by appointment
- Location: Jeff 220

TA INTRODUCTION

- Office Hours: 3: 00 5: 00 PM on Tuesday
- Office Location: Room 102 B, Main Geography Building

HOW ABOUT YOU?

Your background (e.g., name, major, where you come from)

What is your funniest thing that happened during your winter break?

What relevant experience do you have with statistics?

What are your expectations for this course?

RESOURCES

- Introductory text is available for free via <u>https://www.openintro.org/book/os/</u>
- Intermediate book is Joseph Hair, William Black, Barry Babin and Rolph Anderson. Multivariate Data Analysis.
 Edition 7 or 8. Upper Saddle River NJ: Prentice Hall.

Amazon.com usually offers used copies for less than \$30.

https://www.amazon.com/Multivariate-Analysis-Joseph-Anderson-William/dp/9353501350/ref=pd_cp_14_2/144-5828787-2724822? encoding=UTF8&pd_rd_i=9353501350&pd_rd_r=397e3554-2af1-476b-8336-5c1018af6453&pd_rd_w=Q3cmn&pd_rd_wg=XzEv7&pf_rd_p=0e5324e1-c848-4872-bbd5-5be6baedf80e&pf_rd_r=FYB3ZG6A42Z1ANQMRH1V&psc=1&refRID=FYB3ZG6A42Z1ANQMRH1V

COURSE REQUIREMENTS

Assignments: 9 in total

For any graded assignment, if the you do not agree with the grade received, the instructor and TA must be notified within one week after the assignment is graded.

- Late policy for lab
- One final project (oral presentation and paper report)

EVALUATION

- Assignments
 80% = nine assignments
 - Each assignment need to be completed using R
 - > You can ask for help with assignments from the instructor and TA, No plagiarism is allowed
- Final Project 20% = 15% oral + 5% written
 - > To apply & to interpret statistical procedures
 - > To make an oral presentation of a statistical analysis
 - > To write a report

A	93.0 - 100.0	B+	88.0 - 89.9	C+	77.0 - 79.9	D+	67.0 - 69.9
		В	83.0 - 87.9	С	73.0 - 76.9	D	60.0 - 66.9
A -	90.0 - 92.9	B-	80.0 - 82.9	C-	70.0 - 72.9	F	0.0 - 59.9

IMPORTANT DATES

- Jan 22. Add/Drop ends (& last day to request audit) Full Semester
- No class
 - > Jan 20. University holiday
 - Feb 17. Wellness day
 - March 3-7 Spring Break
 - March 24-27 AAG Conference
- Final project
 - > April 14-24 Working on final project
 - > April 28 May 1 Final project presentation
 - > May 5 Final report due

COURSE WEBSITES ON GITHUB



Spatial Database



Python Programming



Intermediate Statistics



Web Mapping



GIS HELP DESK

NEW SPRING 2025 HOURS:

Need help with GIS? Can't make it to visit your professor or TA for office hours? Stop by our GIS Help Desk in the Jefferson Tower (6th floor), or in Jefferson 220A (2nd floor Jefferson building, Geog Main office) on Wednesdays, and visit one of our Help Desk Assistants!

MONDAYS	SASHA 9 AM - 11 AM
TUESDAYS	SASHA 9 AM - 12 PM
WEDNESDAYS @ JF220A	SASHA 9 AM - 12 PM
	WYNNIE 2 PM - 5 PM
THURSDAYS	WYNNIE 2 PM - 5 PM
FRIDAYS	WYNNIE 3 PM - 5 PM

GIS Help Desk hours follow a University schedule. If the University is closed from remote/online operations for any reason, the GIS Help Desk may also be unavailable. GIS Help Desk Assistants operate on a first-come, first-serve drop-in basis. Contact Marjorie Miller (marmiller@clarku.edu) with any questions, or call the Geography Main Office at 508-793-7336 for more information. The GIS Help Desk is sponsored by the Graduate School of Geography at Clark University.

SPRING 2025 HOURS

GIS HELP DESK

Please contact either of our GIS Help Desk Assistants during their specified hours for more information.

SASHA GANNON | GEOG '24 MS-GIS '25



MONDAYS 9AM – 11AM TUESDAYS 9AM – 12PM WEDNESDAYS* 9AM –12PM



WYNNIE GROSS | GEOG '24 MS-GIS '25

WEDNESDAYS* 3PM – 5PM THURSDAYS 2PM – 5PM FRIDAYS 3PM – 5PM

Help Desk hours follow a University schedule. If the University is closed (or if the Geography office is closed), the GIS Help Desk will be unavailable. Hours may be limited or extended during midterms/final exams. Any changes will be announced or posted on the Clark University Geography Facebook page. Appointments operate on a first-come, first-serve drop-in basis unless otherwise scheduled.
 508.793.7336 JEFFERSON TOWER, 6TH FLOOR; *WEDNESDAYS ARE IN JEFFERSON 220A, GEOG MAIN OFFICE, JEFFERSON BUILDING, 2ND FLOOR*

WEEKLY SCHEDULE

Instructor: Yanan Wu TA: Khadija Nisar

Spring 2025



Download

R Project

What's New? Reporting Bugs

Conferences

Get Involved: Mailing Lists

Get Involved: Contributing Developer Pages

Search

R Blog

About R

Logo Contributors

CRAN

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To **download** R, please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.

News

- The useR! 2025 conference will take place at Duke University, in Durham, NC, USA, August 8-10.
- R version 4.4.2 (Pile of Leaves) has been released on 2024-10-31.
- We are deeply sorry to announce that our friend and colleague Friedrich (Fritz) Leisch has died. Read our tribute to Fritz here.
- R version 4.3.3 (Angel Food Cake) (wrap-up of 4.3.x) was released on 2024-02-29.
- You can support the R Foundation with a renewable subscription as a supporting member.



- COURSE INTRODUCTION
 SOFTWARE INSTALLATION
- R programming
- R-Studio
 - Open-source IDE (integrated development environment)

WEEK 02 & 03

Types of Statistics Descriptive Vs Inferential



Steps for Performing Exploratory Data Analysis

æ





IMAGE SOURCE: <u>HTTPS://WWW.LINKEDIN.COM/PULSE/DESCRIPTIVE-STATISTICS-VS-INFERENTIAL-MEASURE-CENTRAL-NHTOC/</u>

HTTPS://WWW.GEEKSFORGEEKS.ORG/WHAT-IS-EXPLORATORY-DATA-ANALYSIS/



- Bivariate Regression
 - Key Assumption
 - Slope and Intercept
 - R square
 - Confidence Interval

WEEK 05 & 06

MULTIPLE REGRESSION

3D Regression



REGRESSION CRITICISM





Happy Spring Break!







WEEK 09 ANOVA

- Assumption of ANOVA
- Type of ANOVA
- Diagnostic in ANOVA

LOGISTIC REGRESSION





2025 AAG Conference



Community 🗸

🖌 🛛 About Us 🖌

Professional Journey 💊

Where Geography Comes Together

ABOUT US

 Principal Component Analysis

A powerful dimensionality reduction technique used in statistics and machine learning to simplify complex datasets.





GEOGRAPHICALLY WEIGHTED REGRESSION



Spatial Autocorrelation

PREAMBLE

- A good online textbook, <u>Hands-on Programming with R</u>, for R beginner.
- Explore the R project website: <u>https://www.r-project.org/</u>
- Explore R Studio: <u>https://posit.co/</u>

R-INTRODUCTION

An online free learning source: <u>An Introduction to R</u>



Data Manipulation

Data Visualization

- Data Visualization Section in <u>R for Data Science</u>
- Modern Data Visualization with R

Interactive Applications

Shiny Gallery in R

R - INTRODUCTION

Statistical Analysis

- Descriptive analysis (mean, median., etc)
- Regression analyses (linear, logistic, ect)
- Time series analysis (ARIMA, etc)
- Multivariate analysis (PCA, factor analysis)
- A handbook of statistical analysis in R

Geospatial Data Analysis

- Handle raster and vector data
- Analyze spatial data with sf, sp or raster

R - INTRODUCTION

Machine Learning

- Implement supervised learning (classification, regression).
- Apply unsupervised learning (clustering, dimensionality reduction).
- Perform deep learning with packages like keras or torch.
- Evaluate models using cross-validation and other metrics.

LAB SESSION

Instructor: Yanan Wu TA: Khadija Nisar

Spring 2025

OVERVIEW OF RSTUDIO INTERFACE

- The panes
 - Left pane: R console
 - Right top pane: includes tabs such as *Environment* and *History*
 - Right bottom pane: File, Plots, Packages, Help and Viewer

🚰 🔹 📄 🛃 🥌 🌾 Go to file/function 🛛 🖾 🔹 Addins 🔹			Projec
Terminal × Background Jobs ×	Environment History Connections Tutorial		
	🧭 🦙 🖬 🖙 Import Dataset 🔹 🌗 144 MB 🔹 🏒		📃 List
	R = 🛛 👼 Giobal Environment =		Q,
	Environment is	: empty	
	Car Back Behave Man Many Republic		
	Files Plots Packages Help Weave Presentation		
	Files Plots Packages Help Viewer Presentation Quite viscer Soler () Nere Bark File * O Conte a, Barane @ More * ↑ Nore		
	Files Plots Packages Help Wewer Presentation QL New Force O New Bars File • O Dete:	520	Modified
	Files Plots Packages Help Viewer Presentation QL Nork Society Stoce Bank File • O Detter (a) Storane (a) More • A Name A Name (a) Storane (a) Storane Image: Storane (a) Storane (a) Storane	5ae 76.68	Modified 5ep 29, 2024, 12:00
	Files Piels Packages Help Viewer Presentation Q New Forder O New Sank File O Detete Iterame Iterame Iterame A lane O Rota Image: Strate Iterame Iterame	500 7.6 KB 6.8 KB	Modified Sep 29, 2024, 12:00 Cot 13, 2024, 8:59 P
	Files Plots Packages Help Viewer Presentation Q Ison Bank File + ○ Center (a) Baname On the aname ↓ Name ○ A loane ○ Romos ○ Romos ○ A loane ○	Ster 7.6 40 6.8 108	Modified Sep 29, 2024, 12:00 Oct 13, 2024, 8:59 P
	File Plots Packages Help Viewer Presentation Q1 here in Scient (Q) Scient Sank File * Q berter Scient Scient (Q) Q1 here A hore Q2 here A hore Q2 here A hore Q3 hore A hore Q4 hore A hore	55a 7.5 KB 6.8 KB	Modified 59 29, 2034, 12:00 Oct 13, 2024, 8:59 P
	Files Plots Pachages Holp Wener Presentation ● Neter Soler ● Neter Soler ● Decise ▲ Name ● ● Nato ● ● ● ● Nato ● ● Arcois ● ■ Arcois ● ■ ● ■ Arcois ● ● ■ Arcois ● ● ■ Arcois ● ● ■ Arcois ●	50e 7.6 48 6.8 45	Modified Sep 29, 2024, 12:00 Oct 13, 2024, 8:59 P
	Fets Pots Packages Help Viewer Presentation Q1 keek solder Q1 keek Bask Fite * Q bete a, Banane Q0 kore * Q1 keek solder A kane Q1 keek solder A kane Q1 keek solder A kane Q1 keek solder A kane Q1 keek solder A kane Q1 keek solder A kane Q2 korder A kane Q2 keek solder Q2 keek	52# 76.40 68.48 24.40	Modified 5ep 29, 9024, 12:00 Oct 13, 2024, 8:59 P Dec 9, 2024, 12:24 P
	Files Picks Picks Picks Picks Q New Factor O Desk Desk Q New Factor O Desk Desk Q None O Desk Desk Q None O Desk Desk Q None O Desk Desk Q Arcidition Desk Desk Q Desk Desk Desk	512e 7.5 KB 6.8 KB 2.4 KB 407 5	Modified 5ep 29, 2024, 12:00 Oct 13, 2024, 8:59 P Dec 9, 2024, 12:24 P
	Files Plots Packages Help Viewer Presentation Q1 Nork Noder Q1 Nork Bank Title * Q Detter (a) Benare (a) More * A None A None A None A None (a) Anota Q2 Restory (a) Anota (a) Anota (a) Anota Q2 Anota (a) Anota (a) Anota (a) Anota Q3 Detted S Po 3.3 (a) Custom Office Emplates (b) Detted S Po 1.3 Q2 Detted S Po 1.4 (b) Detted S Po 1.5 (b) Detted S Po 1.5 Q3 Detted S Po 1.5 (b) Detted S Po 1.5 (c) Detted S Po 1.5 Q4 Detted S Po 1.5 (c) Detted S Po 1.5 (c) Detted S Po 1.5 Q4 Detted S Po 1.5 (c) Detted S Po 1.5 (c) Detted S Po 1.5 Q4 Detted S Po 1.5 (c) Detted S Po 1.5 (c) Detted S Po 1.5 Q5 Detted S Po 1.5 (c) Detted S Po 1.5 (c) Detted S Po 1.5 Q5 Detted S Po 1.5 (c) Detted S Po 1.5 (c) Detted S Po 1.5	528 7.6 % 6.8 % 2.4 % 402 % 205 %	Modified 5ep 29, 2024, 12:00 Oct 13, 2024, 8:59 P Dec 9, 2024, 12:24 F 5ep 12, 2024, 10:24 F 5ep 12, 2024, 10:24
	Files Plots Packages Holp Wear Presentation ● More ● Boats ● ● Boats Boats	500 7.6 KB 6.8 KB 2.4 KB 402 B 2040 2 KB 2040 2 KB 2040 2 KB	Modified Sep 39, 3024, 1200 Oct 13, 2024, 1204 Sep 12, 2024, 1204 Sep 12, 2024, 1204 New 27, 2024, 849 New 27, 2024, 849
	Fals Packages Help Viewer Presentation Image: Note Noder Note Bank File Optice State Image: Note Noder Note Noder Didite Image: Noder Image: Note Noder Note Noder Note Noder Image: Noder Image: Noder Noder Note Noder Noder Image: Noder Image: Noder Noder Noder Noder Noder Noder Noder Noder Image: Noder N	528 76 KB 68 KB 422 B 236 2 KB 236 2 KB	Mooffred Sep 29, 3024, 1200 Oct 13, 2024, 859 P Sep 21, 2024, 1020 Sep 12, 2024, 1020 Horv 27, 2024, 464 Horv 27, 2024, 4649

OVERVIEW OF RSTUDIO INTERFACE

- Starts a new pane on the left
 - File New File R Script

🔹 🥸 💣 📲 🔚 🛑 🍺 Go to file/function 🔡 🗄 🔹 Addins 👻				Project: I
Untitled1 ×	-0	Environment History Connections Tutorial		
🗇 🗇 🗐 🔚 🗆 Source on Save 🔍 🎢 🚽 🗐	🖚 Run 🏞 🏠 🔄 Source 🗸 🖻	😅 🕞 📑 Import Dataset 👻 🌗 144 MiB 👻 🎻		🗏 List 🝷
1		R 🝷 💁 Global Environment 👻		Q,
		Environment is empty		
11 (Top Leve) : nsole Terminal # Background Jobs //	R Soriet :=			
		Files Plots Packages Help Viewer Presentation		
		New Folder V New Blank File V Delete Rename Wore V		
		A Name	Size	Modified
		RData	7.6 KB	Sep 29, 2024, 12:00
		Rhistory	6.8 KR	Oct 13, 2024, 8:59 P
		Alcois		
		ArrGIS Dep 2.2		
		ArcGIS Pro 3.3		
		ArcGIS Pro 3.3 Giscon Confice Templates	2.4.15	
		ArcGIS Pro 3.3 Custom Office Templates Defaultridg	2.4 KB	Dec 9, 2024, 12:24 F
		Arctis Fro 3.3 Gottom Office Templates Custom Office Templates Oddunturdap Oddunturdap Oddunturdap	2.4 KB 402 B	Dec 9, 2024, 12:24 P Sep 12, 2024, 10:03
		ArcGIS Pro 3.3 Custom Office Templates Default.rdp D	2.4 KB 402 B 236.2 KB	Dec 9, 2024, 12:24 P Sep 12, 2024, 10:03 Nov 27, 2024, 8:48 P
		ArctiS Pro 3.3 ArctiS Pro 3.3 ArctiS Pro 3.3 Bernet Custom Office Templates Order Strategy Order Strategy Order Strategy Order Strategy Dr. Vanan Wu RF02.pdf	2.4 KB 402 B 236.2 KB 236.2 KB	Dec 9, 2024, 12:24 i Sep 12, 2024, 10:03 Nov 27, 2024, 8:48 Nov 27, 2024, 8:49
		Arctil S Pro 3.3 Custom Office Templates Custom Office Templates Office Temp	2.4 KB 402 B 236.2 KB 236.2 KB	Dec 9, 2024, 12:24 f Sep 12, 2024, 10:03 Nov 27, 2024, 8:48 f Nov 27, 2024, 8:49 f

OVERVIEW OF KEYBOARD SHORTCUTS

- Keyboard shortcuts
 - Help Keyboard Shortcuts Help



OVERVIEW OF CHEATSHEETS

- Cheatsheets in Rstudio
 - Help Cheatsheets

s	Help					
•	R Help					
	Search R Help	Alt+Ctrl+F1				
	About RStudio			📑 Run 🍠	Ŷ	Ð
	Check for Updates					
	Accessibility		۲			
	RStudio Docs					
	RStudio Community Forum					
	Cheat Sheets		•	RStudio IDE Cheat Sheet		
	Keyboard Shortcuts Help	Alt+Shift+K		Data Transformation with dplyr		
	Markdown Quick Reference			Data Visualization with ggplot2		
	Roxygen Quick Reference			List manipulation with purrr		
	Diagnostics		Þ	Package Development with devtools		
				Web Applications with shiny		
				Interfacing Spark with sparklyr		
				R Markdown Cheat Sheet		
			_	R Markdown Reference Guide	_	
				Browse Cheat Sheets		

GLOBAL SETTING

- .RData
 - Save your workspace, including variables, data frames, lists, and other objects
- Cons
 - Causing confusions especially when we share code with others and assume they have this .Rdata file
- Tools Global Options
 - Change the setting as below

Workspace

Restore .RData into workspace at startup:

Save workspace to .RData on exit: Never



WORKING DIRECTORY

- Get working directory where your scripts and workspaces are stored
 - getwd()
 - Run this command:
 - Ctrl + Enter
 - Or Run in R
 - The returning strings, e.g., "C:/Users/yy00021/Documents" is the path to the working directory
 - The windows convention uses slash \ to separate sub-directories
 - However, R uses forward slash / or a double backward slash \\
- Change working directory
 - I suggest you to setup a specific directory for this course
 - setwd('D:\\Spring2025\\geog247')
 - Now check your working directory again



CONSOLE WINDOW

- The character > in CONSOLE window indicates that R is ready to receive new commands
- It show up when R completed executing a script



TERMINATE SCRIPT

The Esc Key or pressing on the CONSOLE window to terminate the script

```
## Terminate script
i <- 1
while (i>0) {
  print('good')
}
```

GET HELP

- Get help for activate libraries
 - help('dplyr')
 - ?dplyr
- Get help for all installed libraries
 - ??dplyr
 - help.search('dplyr')

INTERACTING WITH THE R-CONSOLE

- All commands (or programs) can be stored in external *.R script-files
- Single command or a set of highlighted commands can be run using shortcut (shift+enter) or Run button in R
- All commands can be run use the Source button in R
- Scroll through the history of previously commands in R
- Using shortcut key (Ctrl + L) or broom icon to clean the Console window

VARIABLES IN R

- Variable names
 - Variables are created using the assignment operator
 - Variables can store different types of data (numeric, character, logical, etc.).
 - Variables can be reassigned new values anytime.
 - The <u>document</u> shows professional naming for your code
- Object in the ENVIRONMENT
 - Any data structure or function that is defined using commands becomes an object in the ENVIRONMENT



- Remove objects
 - The objects can be removed from the ENVIRONMENT
 - rm(x)
- Clean ENVIRONMENT
 - Broom icon in the ENVIRONMENT mean bar
 - or rm(list=ls())

LIST IN R

- Creating a list
 - A list in R is a flexible data structure that can contain elements of different types: numbers, characters, vectors, matrices, data frames, or even other lists.
 - It's like a container for multiple objects.
- Accessing elements in a list
 - Use [[]] to access elements by position or name.
 - Use \$ to access elements by name.

DATA SETS

- Read csv
 - read.csv() for reading CSV files.
- Check columns
 - Accessing column names using colnames()
- Add new columns
 - Adding columns based on calculations or conditions

PRACTICES

Instructor: Yanan Wu TA: Khadija Nisar

Spring 2025

PRACTICES

- Explore Tools and Help in RStudio
- Explore the different tables in RStudio