

Title: Fundamentals of R and Posit (R Studio) – A One Day Short Course

Description: This one-day intensive course is designed to introduce participants to R and RStudio (Posit), equipping them with the fundamental skills required to start their journey in data analysis and programming. The course begins with an overview of R and RStudio's environment, highlighting how these tools are used for data manipulation, analysis, and visualization in various fields. Attendees will learn how to import, clean, and manipulate datasets using R's **dplyr** package, along with how to create data visualizations using R's **ggplot2** package. Throughout, this course focuses on the principles of data transparency and reproducibility. By the end of the day, participants will have a solid foundation in R and RStudio, enabling them to confidently apply their new skills to real-world data analysis projects. This course is ideal for beginners with no prior experience in R or those looking to refresh their skills.

Requirements – A Laptop computer running any of Linux, Mac or Windows OS

Prior to class:

- Install R following directions from here: <https://cran.r-project.org/>
- Install Rstudio (Posit) following directions from here: <https://posit.co/downloads/>

1. Concepts – Subject to Change

- Rstudio IDE overview
 - Installing and loading packages
 - Opening and closing scripts
 - Help documentation access
 - The working directory and relative paths
 - Assignment operator and object browser
 - Rstudio Script Types – this course will focus on R markdown
 - .R vs .RMD and knitting to HTML output
- Importing and exporting data
 - Base R functions for delimited files
 - data.table package and fread()
 - Excel files from readxl package
- Working with Vectors
 - Vector Creation and Types of Vectors (character, numerical, logical) using c()
 - Vector Operations (sum(), mean(), length(), min(), max())
 - Logical Operations
 - Indexing and Subsetting
- Working with Data Frames
 - Creation of data.frames()
 - Inspection with head(), tail(), dim(), str(), nrow(), ncol()
 - The \$ notation
 - Adding columns and rows with rbind() and cbind()
 - Column calculations
- Data Munging with dplyr()

- Filtering Rows and Columns with `filter()` and `select()`
 - Creation of new variables with `mutate()`
 - Summarizing data with `summarise()`
 - Sorting and grouping operations with `arrange()` and `group_by()`
- Merging datasets with `dplyr()`
 - Left, right, full, inner joins
- Visualization
 - Base R graphics
 - Graphics in `ggplot2`
- Functional programming
 - Parameters
 - Default values
 - Loops

2. **Agenda** (subject to change)

Start	End	Program
8:00 AM	8:30 AM	<i>Arrival</i>
8:30 AM	9:30 AM	R and Rstudio overview, navigation and setup
9:30 AM	10:00 AM	Getting data in and out of R
10:00 AM	10:15 AM	<i>Short Break</i>
10:15 AM	11:00 AM	Working with vectors
11:00 AM	12:00 PM	Working with data frames
12:00 PM	12:45 PM	<i>Lunch Break</i>
12:45 PM	1:45 PM	Data munging with <i>dplyr</i>
1:45 PM	2:30 PM	Merging data in R with <i>dplyr</i>
2:30 PM	3:00 PM	Data visualization I
3:00 PM	3:15 PM	<i>Short Break</i>

3:15 PM	4:15 PM	Data visualization II
4:15 PM	5:00 PM	Intro to functional programming

3. Instructor

John S. House Ph.D., M.Stat., Staff Scientist – NIEHS

Adjunct Associate Professor – NCSU

4. Location

Talley 4280, NCSU Main Campus, March 30, 2024, 8:00 AM – 5:00 PM.

5. Extra Information

Requirements – A Laptop computer running any of Linux, Mac or Windows OS.

Prior to class:

- Install R following directions from here: <https://cran.r-project.org/>
- Install Rstudio (Posit) following directions from here: <https://posit.co/downloads/>